THE LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE

Essays in Applied Macroeconomics

Federico Rossi

A thesis submitted to the Department of Economics of the London School of Economics for the degree of Doctor of Philosophy.

London, July 2017

Declaration

I certify that the thesis I have presented for examination for the PhD degree of the London School of Economics and Political Science is solely my own work other than where I have clearly indicated that it is the work of others (in which case the extent of any work carried out jointly by me and any other person is clearly identified in it).

I warrant that this authorisation does not, to the best of my belief, infringe the rights of any third party.

I declare that my thesis consists of approximately 50,000 words.

Statement of conjoint work

I certify that Chapter 1 was co-authored with Marta De Philippis, and that I contributed 50% of the work.

Acknowledgements

I wish to express my gratitude to several people that accompanied me through this journey. My supervisor, Francesco Caselli, was an incredible source of insights and inspiration. He taught me to think boldly, to question all my ideas, and to believe in them nevertheless. I am also grateful to the rest of the faculty at the LSE, and in particular to Silvana Tenreyro and Wouter den Haan for their invaluable support.

I wish to thank the UniCredit & Universities Foundation, and in particular Annalisa Aleati and Giannantonio De Roni, for their financial support during my MRes and countless enjoyable events in London, Milan and Sicily. I am also grateful to the CEP, the Department of Economics and the CFM for the stimulating job opportunities they provided me with at different stages of my PhD.

The greatest legacy of these years at the LSE is undoubtedly the tight bonds I created with several friends and colleagues. I am grateful to Francesco and Nicola for being part of the most solid trio ever, to Diego and Soledad for the Italian-Argentinian connection and to Alexia, Davide, Gianpaolo, Jiajia, Junichi, Kilian, Oriol, Matteo, Stephan, Thomas and many others for having shared some the most fun moments in the last few years. It has also been a great privilege to collaborate with Marta on the first chapter of this thesis.

I wish to thank my family, Giulia, Grazia and Villiam, for their patience and unconditional support over these years and before. I know that Ilario would have proudly stored a copy of this thesis, together with the previous ones, and I owe a special thanks to Liana and him for having always believed in me.

Finally, my deepest gratitude goes to Maddi. She is the person that more than anyone else has shared with me this adventure, and the one with whom I want to share all future ones.

Abstract

This thesis consists of three chapters, examining the interrelation between human capital and country-level outcomes from different perspectives.

Chapter 1, co-authored with Marta De Philippis, studies the contribution of parental influence to cross-country gaps in human capital performance. We compare the school performance of second-generation immigrants from different nationalities but educated in the same school, and find that those whose parents come from high-scoring countries in international standardized tests do better than their peers. The gap is larger when parents have little education and have recently emigrated, suggesting the importance of country-specific cultural traits that parents progressively lose as they integrate in the new host countries. Parental influence accounts for between 14% and 20% of the cross-country variance in test scores.

Chapter 2 studies the macroeconomic consequences of the inequality of educational opportunities. I discuss how family income shapes college opportunities for US students, even when its correlation with academic ability is taken into account. I propose a general equilibrium model to estimate the productivity losses deriving from the fact that human capital investment is not always allocated where its marginal product would be highest. Using the equilibrium conditions of the model, I back out the value of barriers to college investment for disadvantaged students from data on family income, ability, schooling and wages. Counterfactual experiments suggest that a more meritocratic access to college education could boost output by approximately 11%, and wages by between 9% and 12%. I conclude that returns from policies aimed to expand college opportunities are potentially very large.

Chapter 3 studies how the relative productivity of skilled and unskilled labor varies across countries. I use both micro-data for countries at different stages of development and other sources to document that the skill premium varies little between rich and poor countries, in spite of large differences in the relative skill supply. This pattern is consistent with the view that the relative productivity of skilled workers is higher in rich countries. I propose a methodology based on the comparison of labor market outcomes of immigrants with different levels of educational attainment to discriminate between technology and unobserved human capital as drivers of these patterns. I find that human capital quality plays a minor role in explaining cross-country differences in relative skill efficiency.

Contents

List of Tables

List of Figures

Chapter 1

Parents, Schools and Human Capital Differences across Countries

1.1 Introduction

Human capital varies greatly across countries, in terms of both years of schooling (Barro and Lee, 2013) and results in international standardized tests. East Asian countries consistently position themselves at the top of international test rankings, while several Southern European and Latin American countries perform poorly. An emerging strand of the growth literature argues that human capital accounts for a substantial part of cross-country differences in economic performance (Schoellman, 2012; Jones, 2014; Lagakos et al., 2016), especially when measured by standardized tests (Hanushek and Woessmann, 2012a).

Given the role that gaps in human capital measures play in the academic and policy debates, it is important to understand where they come from. Most of the discussion on standardized tests relies on (and argues in favor of) an interpretation of the results as measures of school quality.[1] More broadly, the literature on cross-country differences in educational attainment emphasizes country-specific factors such as access to public education, sectoral composition and skill premia that shape the costs and expected benefits of human capital investments. On the other hand, studies on skill formation at the individual level argue that parents and the home environment are of great importance (Almond and Currie, 2011). A natural question then is whether variation in these factors is relevant also at the country level.[2]

In this paper we investigate how much of the cross-country variation in test scores can be attributed to differences in broadly defined parental influence, and what the nature of these differences is in the first place. The analysis involves difficult challenges. Parental influence is generally unobservable, and, even when proxies are available, cross-country comparisons cannot separately identify its contribution from the one of school quality or other institutional

[1]The popular press is rich of anecdotes about the severity of school curricula in East Asian countries. For a recent example, see Jeevan Vasagar, "Why Singapore's kids are so good at maths", *Financial Times*, July 22, 2016.

[2]Anecdotal evidence suggests that indeed parenting styles and parental attitudes towards education vary across countries; for example, the international bestseller by Chua (2011) coined the expression "Tiger Mother" to describe demanding Asian mothers, focusing on their children's academic excellence.

factors. We overcome these difficulties by adopting an indirect approach, based on the analysis of second-generation immigrants. We compare the performance of students born and educated in a given country and, for part of the analysis, in the same school, but with parents of different nationalities. Since factors such as the educational curriculum, teachers and the institutional setting (as well as other individual-level characteristics) are kept fixed in this comparison, we argue that we can reasonably attribute any residual difference to differential influences exerted by parents. We then use the results from this empirical exercise to decompose the cross-country variation in test scores between different sources, shedding light on the nature of these gaps.

Our results point towards a substantial role for parents. First, we document that the PISA performance of second-generation immigrant pupils, living in the same country and studying in the same school, is closely related to the one of natives from the country of origin of their parents: the best performing second-generation immigrants are those whose parents come from countries where natives are particularly successful in standardized tests.[3] This holds true when controlling for parental education, socio-economic status and other characteristics of parents' countries of origin. Moreover, we find a similar result for a different schooling outcome in a specific host country, which is grade repetition in the United States. These findings do not appear to be driven by a pattern of differential selection into emigration.

We construct a country-specific measure of parental influence, which combines the estimated effect of average observable and unobservable (as captured by country of origin fixed effects) parental characteristics. According to our estimates, between 14% and 20% of the total cross-country variation in test scores can be accounted for by differences in this term. Parental influence is responsible for a substantial share of the East Asian out-performance: on average, between 22% and 58% of the gap between Chinese and non-Chinese native students is persistent across second-generation immigrants.

We then focus on the US data to explore the nature of these differences in terms of parental influence. We show that the relationship between the performance of second-generation immigrants and the average score in the parents' country of origin is strongest for parents with little or no formal education. This suggests that our results are not driven by the quality of education received by parents in their home country. Moreover, the relationship weakens if parents have spent more years in the host country, suggesting the importance of country-specific "cultural" traits, that are progressively lost by emigrants as they integrate in their new host country. This interpretation is reinforced by the fact that part of the variation in second-generation immigrants' performance is accounted for by proxies for cultural traits likely to be conducive to human capital investment, such as long-term orientation, locus of control and attitudes towards leisure. Finally, time use data for immigrants in the US show that parents from high PISA countries spend more time on various forms of child care.

This paper contributes to the debate on cross-country differences in human capital. Several

[3]Throughout the paper, we call natives those students born in the country where they are taking the test and whose parents are born in the same country as well. On average, across countries participating to the PISA test, natives represent 78% of the target population. Students born in a country different from the one where they are taking the test are excluded from the analysis.

papers study the importance of characteristics of the school system (Woessmann, 2016), while other contributions focus on country-specific factors that shape human capital investment (Bils and Klenow, 2000; Manuelli and Seshadri, 2014; Schoellman, 2016). Our emphasis on parents is shared by Doepke and Zilibotti (2017), who develop a model of preference transmission to explain the international variation in parenting styles as a function of local economic conditions. We contribute to this literature by quantifying and characterizing cross-country differences in parental influence, in a setting where other country-specific factors are arguably not operative.

We also relate to a wide literature across economics and sociology on the school performance of first- and second-generation immigrant children (see Levels et al., 2008; Dustmann et al., 2012, for broad reviews). Differently from these papers, our objective is to understand gaps in performance between natives of different nationalities, and our focus on second-generation immigrants is mostly instrumental in that it provides us with an empirical strategy to discriminate between possible sources for these gaps.[4] In addition, we conduct our analysis on a broad sample of host and origin countries (while, for example, Dustmann et al. (2012) focus on Turkish immigrants, and Jerrim (2015) on East Asian immigrants), and we rely on several additional sources to provide evidence on the mechanisms underlying our results.

Our paper shares the approach of a large literature that looks at first- and second-generation immigrants to identify the importance of "portable" cultural traits for various outcomes (the so-called "epidemiological approach"; see among others Giuliano, 2007; Fernandez and Fogli, 2009; Fernandez, 2011). Differently from these papers, we study the school performance of the second generation, and use the results to quantify the importance of parents for cross-country differences in the same outcome. While most of the focus in this literature is on immigrants in the US, our sample includes a large set of both host and source countries, allowing us to exploit variation in both dimensions.[5]

The paper is structured as follows. Section 1.2 discusses different forms of parental influence, and clarifies which our empirical approach can capture. Section 1.3 describes the data, while Section 1.4 shows evidence on the performance of second-generation immigrants. Section 3.4 addresses selection. Section 1.6 quantifies the importance of parental influence, while Section 1.7 explores the mechanisms behind our results. Finally, Section 3.5 concludes.

[4]Most papers have focused on the comparison between immigrants and natives in the host country. Like us, Levels et al. (2008) and Dronkers and de Heus (2016) compare the performance of (a combination of) first- and second-generation immigrants across countries of origin. However, they do not relate those to the performances of natives in the countries of origin, nor explore the implications in terms of cross-country gaps in performance. Yet another distinct strategy is the one in Borjas (1992), who relates the average educational attainment of ethnic groups residing in the US (what he calls "ethnic capital") to schooling and wages of the following generation. We discuss this and other channels through which immigrant parents' ethnic network might affect children's human capital accumulation in Appendix A.4.

[5]In more recent and independent work, Figlio et al. (2016) adopt a similar methodology to study the effect of long-term orientation on educational performance. Compared to their paper, we do not restrict attention to a specific cultural trait, but study and quantify the overall importance of observables and unobservables parental characteristics for the cross-country variation in human capital achievement. In Section 1.7 we do look explicitly at long-term orientation, among other cultural traits, and confirm the Figlio et al. (2016)'s result that it affects students' performance, even though it cannot account for the whole cross-country variation in parental influence.

1.2 Parental Influence: Definition and Discussion

Parental influence on children's human capital can manifest itself through a number of channels. The activities that parents do with their children (or push them to do on their own), the teachings they pass them and the resources they provide them with all plausibly affect their human capital development. Parents shape children's attitudes towards education and effort, and might have an indirect influence through the example they provide. The genetic transmission of traits that affect learning ability and preferences could also play a role.

Our measure of parental influence, based on school performance gaps across second-generation immigrants, includes the effect of all inputs listed above. While Section 1.7 speaks to the relative importance of some of these channels, the extent to which we can discriminate between them is limited by the fact that most of these factors are difficult to measure.

An important qualification concerns the *reason* behind the supply of different levels of parental inputs. Parents' choices are partially driven by *context-specific* incentives: for example, higher expected returns to skills in the labour market might induce parents to stress the importance of education and hard work (Doepke and Zilibotti, 2017). On the other hand, factors that are *embedded* into parents, independently of the context-specific incentives they face, are also likely to be important: for example, preferences on education and parental productivity in the process of skill transmission fall into this category.

What do we pick up by comparing second-generation immigrants? Institutional factors and the educational system are kept fixed, allowing us to focus on parental influences. In addition, parents in our sample experience similar *context-specific* incentives, since their children face the same educational system and, ruling out differential intentions in terms of future relocation, labor markets with similar characteristics. The relevant source of variation is represented by *embedded* factors, which might differ across parents because of cultural traits or skills inherited from their country of origin.

Taking stock of this discussion, our methodology allows us to isolate the importance for cross-country differences in human capital of inputs driven by factors *embedded* into parents. This is an interesting dimension for the analysis of cross-country gaps, since factors that lead parents to invest differentially in their children independently from the local economic and social conditions are likely to be very persistent over time, and perhaps particularly hard to affect through policy.

1.3 Data

Our main data come from the 2003, 2006, 2009 and 2012 waves of the PISA test. PISA is a triennial survey of the knowledge and skills of 15-year-old children, explicitly designed to allow comparisons across countries. The test covers three subjects: reading, mathematics and science. We standardize scores to have mean 0 and individual-level standard deviation 1 across all countries (pooled, equally weighted) participating in at least one wave of the test.[6]

Results for all subjects vary greatly across countries. Figure 1.1 shows the average math score of native students for all countries that participated to at least one wave of the PISA test (pooled across all available waves). Chinese students score 1.3 standard deviations higher than the average, and almost 3 standard deviations better than the worst-performing countries.[7] These magnitudes are striking; according to OECD (2012a), a gap of 0.4 on this scale corresponds to what is learned in an average year of schooling. There is substantial geographical clustering: East Asian countries occupy the first positions of the ranking, followed by several Western European countries; Southern European countries concentrate in the middle of the distribution, while Latin American countries are below the average. The superior performance of East Asian students is stronger in mathematics, but the ranking across regions is quite stable across subjects (see Table A.1 in the Appendix for the average scores in these and other broadly defined regions).

A Student Questionnaire provides basic demographic information on students and parents, including their country of birth, education, employment and the ISEI index of socioeconomic status.[8] Our sample includes 40,067 second-generation immigrants on the mother's side and 40,304 on the father's side, from 49 and 48 different countries of origin and distributed across 39 host countries.[9] Sample sizes vary greatly, and for some countries of origin we have only a few parents to work with (see Tables A.2 and A.3 in the Appendix for summary statistics by origin and host country). To account for this, we weight countries of origin by the number of second-generation immigrants in the sample when considering cross-country patterns, and we present country-specific estimates for a "core sample" of 37 countries from which we have at least 100 emigrant parents. Solid bars in Figure 1.1 correspond to countries for which we observe second-generation immigrants, and the black ones identify the "core sample". Descriptive statistics for second-generation immigrants on the mother's side are provided in Panel A of Table 1.1.

Our second source is the US Census. We use the 1% sample for 1970 and 5% sample for 1980. We follow Oreopoulos and Page (2006) in combining information on children's age and grade currently attended to construct an indicator of whether or not students have repeated any grade.[10] We classify a child as a repeater if his or her educational attainment is below the mode for the corresponding state, age, quarter of birth, and census year cell.[11] We focus on children

[6]The results are not presented as point estimates but rather as "plausible values": the OECD estimates for each student a probability distribution of scores, and randomly draws from it five values. Following OECD (2009), we compute variances of all functions of test scores as the average of the 5 variances estimated with each set of plausible values, and standard deviations as the square root of the corresponding quantities.

[7]The PISA test in China is held in Shanghai only, and is not representative of the whole country. In contrast, Chinese-born emigrant parents in our sample might come from any part of the country. This suggests a pattern of selection likely to work against our main result; see Section 3.4 for a detailed discussion.

[8]The ISEI index, developed by Ganzeboom et al. (1992), is a measure of occupational status that assigns to each occupation a score from 16 to 90 combining information on average income and education of full-time employed men across several countries.

[9]Individual countries have flexibility on how to classify parents' country of origin. We construct a set of countries consistently defined over time. See Appendix A.1 for the details.

[10]Current grade is only available until 1980, which prevents us from using more recent years.

[11]This grade-for-age measure induces some misclassification, as, for example, students entering school late will be classified as grade repeaters. As discussed in Cascio (2005), this type of misclassification will lead to some

between the ages of 8 and 15. The final sample includes 53,081 second-generation immigrants on the mother's side and 46,410 on the father's side, from 61 countries of origin. Descriptive statistics are provided in Panel B of Table 1.1.

We use the 2002 to 2013 waves of the ATUS-US Time Use Survey to analyze how immigrant parents spend their time. The survey is administered to one person per household, chosen randomly among all individuals at least 15 years old. We compute the total time (in minutes) spent on child care on the previous day, and, following Aguiar and Hurst (2007), three subcategories that split total child care in educational, recreational and basic activities.

Finally, we rely on several other sources to construct our controls at the level of parents' country of origin. We use real GDP per capita from the PWT, average years of schooling for different demographic groups from Barro and Lee (2013), measures of school quality from Bartik (2008) and various proxies for cultural differences from the World Value Survey.

1.4 Reduced Form Evidence

In this section we examine whether the school performance of second-generation immigrants is related to the one of natives in their parents' country of origin. We focus here on second-generation immigrants on the mother's side only. This is only to simplify the exposition, and in Appendix A.2 we show that our results hold without exception when we look at second-generation immigrants on the father's side or at the whole sample of second-generation immigrants and natives. We present results for the PISA and the US Census samples in turn.

1.4.1 PISA

The left panel of Figure 1.2 plots the average score of second-generation immigrants against the average score of natives in the country of origin of their mother, pooled across all available waves. The relationship is positive and tight. While the cross-country variation in natives' performance reflects a combination of school quality, economic, cultural and institutional factors, the fact that these gaps are largely preserved across second-generation students in other countries suggests that parents might play an important role. Of course, this pattern might be driven by factors unrelated to systematic differences in parental influence across countries. We investigate several potential confounders in our regression analysis.

Let T^m_{icst} denote the PISA math score in year t of child i, studying (and born) in country c and in school s, whose mother was born in country m.[12] We estimate variants of the following specification:

$$T^m_{icst} = \theta_0 + \theta_1 T^m + \theta_2' X^m_{icst} + \theta_{cs} + \theta_t + \varepsilon^m_{icst} \tag{1.1}$$

where T^m is the average score of native students in the mother's country of origin, X^m_{icst} is

attenuation bias in all regressions using the grade repetition proxy as outcome variable.

[12]The results are similar for the reading and science tests (see Appendix A.2). Math is often preferred for international comparisons for the relative easiness of defining and quantifying a common set of expected skills (Hanushek and Woessmann, 2012a).

a vector of individual characteristics of students and parents, θ_{cs} is a host country or school (depending on the specification) fixed effect, θ_t is a PISA wave fixed effect and ε^m_{icst} is an error term. We include in X^m_{icst} various parental characteristics likely to be correlated with human capital investments on children, such as parental education, employment status and, for those who are employed, the ISEI index of occupational status.[13] Moreover, by introducing host country (or school) fixed effects we control for differences in the characteristics of the institutional context (or specific school) students are exposed to. The main coefficient of interest is θ_1, which captures the relationship between a second-generation immigrant's performance and the average score of native students in country m.

Table 1.2 shows our results. The sample is limited to second-generation immigrants on the mother's side, and a dummy is included in all specifications to control for whether the father is also foreign born. Standard errors are clustered at the level of the mother's country of origin, and inflated by the estimated measurement error in test scores.[14]

We proceed by progressively adding controls. Column 1 controls for students' baseline characteristics (gender and age in months), fathers' immigrant status and wave fixed effects only. The correlation of interest is strong and highly significant: a gap of one (individual-level) standard deviation in the average score in the mother's country of origin is reflected in a gap of 66% of a standard deviation among second generation immigrants. The coefficient shrinks when we introduce host country (column 2) and, especially, school (column 3) fixed effects, but is still positive and significant. A comparison between the first two specifications and column 3 suggests that, within the same host country, mothers from high PISA countries might send their children to better schools.[15]

Column 4 adds controls for parental education, with the coefficient of interest being hardly affected. This suggests that the estimate of θ_1 is unlikely to be driven by some unobservable parental skills, since we would expect those to be correlated with parental education, and therefore the inclusion of this last variable to matter a lot for our coefficient of interest. Similarly, the introduction of controls for employment and occupational status in column 5 does not change the coefficient of T^m.[16] The last column of Table 1.2 shows that results are not driven only by the performances of students with East Asian origins, since the coefficient is robust to the exclusion of East Asian mothers.

The right panel of Figure 1.2 displays the main result of this section. After we clean the scores of second-generation immigrants from the effect of observable characteristics, including

[13] Information on parental age and number of siblings is available only for a small set of host countries and waves. Our results are robust to the inclusion of these controls in this sub-sample.

[14] As recommended in OECD (2009), each regression is estimated separately for each set of plausible values, and the sampling variance is computed from the average estimated variance-covariance across these specifications. In addition, standard errors are corrected for the imputation variance, which is proportional to the variance of the estimated coefficients across sets of plausible values. In Appendix A.2 we discuss the details of this procedure, and show that the statistical significance of our results is robust to alternative ways to construct the standard errors.

[15] In Appendix A.3 we show that indeed, after controlling for country fixed effects and the usual observable characteristics, a higher PISA score in the country of origin of the mother is positively correlated with several proxies for school quality.

[16] In Appendix A.2 we show that the results are robust to the inclusion of alternative measures of socio-economic status available in the PISA dataset.

school fixed effects, the relationship between the performance of second-generation immigrants and natives in the mother's country of origin country weakens but is still positive and significant.

1.4.2 US Census

We apply a similar specification as in equation (1.1) on the US Census data. The dependent variable is a dummy which takes value one if a child has never repeated any grade. This outcome, while still related to school performance, captures quite a different dimension compared to the PISA score, given that the variation here comes only from the bottom part of the distribution (more than 80% of the students in the sample has never repeated a grade, as shown in Table 1.1) and from students aged 8-15 (while PISA is administered to 15-year-old students only).

The US Census does not include any information on the school children are attending. To capture some of the differences across educational systems within the US, we control for Commuting Zone fixed effects.[17] Compared to the PISA sample, we can control here for a richer set of family characteristics, such as number of siblings, child's and parents' age and family income, as well as for the number of years passed since the mother has migrated to the US.

Table 1.3 shows our results. Once again, the coefficient on T^m is positive and significant throughout. Commuting zones fixed effects and controls for parental education, mother's years since migration and family income explain about two thirds of the gap in performance between second-generation immigrants from high and low PISA countries. According to column 5, the most complete specification, an increase of a standard deviation in the PISA score of students in the mother's country of origin is associated with a higher probability of not having repeated any grade by 2.8 percentage points (3% over the average). As for the PISA specification, the result is robust to the exclusion of East Asian mothers (column 6).

1.5 Selection

As our analysis relies on emigrant parents to make inference on all parents of a given nationality, a concern is that emigrants are not a random sample of the population, and might be selected on unobservable characteristics (such as skills and preferences for education) that matter for children's school performance.[18]

What type of selection should we worry about? Figure 1.3 displays various possibilities. The solid line represents the selection-free relationship between the score of second-generation immigrants and the one of natives from the parents' country of origin, i.e. the relationship that we would be able to observe in a world where emigrant parents were randomly selected from the population. The dashed line represents instead what we would observe in our data under different patterns of selection into emigration. Our parameter of interest is the slope of the solid

[17]Commuting Zones are constructed following Autor and Dorn (2013).

[18]The discussion in this session focuses on selection into emigration. A distinct issue is, conditional on emigration, selection into host countries where parents have different prospects of assimilation. In Appendix A.3 we show that controlling for measures of linguistic and cultural distance between the origin and the host country does not affect our results.

line, or, more generally, the extent to which the relative performance of natives is reflected in the relative performance of second generation immigrants with "representative" parents in terms of unobservable characteristics.

The first panel depicts the case where the extent of selection into emigration (as measured by the gap between the two lines) is the same across countries of origin with different PISA scores. In this case only the estimated intercept is biased, while the inferred slope is not affected. In the second panel we have the case where parents emigrated from countries with high PISA scores are more positively selected than parents emigrated from countries with low PISA scores, while in the third panel we have the opposite case. These patterns of differential selection would lead to a biased estimate of our coefficient of interest, and in particular the case depicted in the second panel could rationalize the findings of the previous sections.

To understand which case is relevant in our setting, we look at differential selection in terms of parental education. While the main threat to our approach is differential selection on unobservables, it seems plausible that several unobservable parental traits that positively affect children's school performance (such as skills and attitudes towards schooling) are positively correlated with parents' own educational achievements. We can therefore alleviate the concerns on differential selection if we can show that the relative "quality" of emigrants compared to stayers is not higher for high PISA countries.[19]

We construct for each parent a measure of selection by computing the difference between his or her years of schooling and the average years of schooling of non-emigrant parents from the same country, and dividing this quantity by the country of origin-specific standard deviation.[20] Figure 1.4 plots the average of this measure across mothers' countries of origin against the average score of native students in those countries. For a majority of countries of origin emigrant mothers are positively selected (that is, our measure is greater than 0), a finding consistent with most of the recent literature (for example, Feliciano (2005b) documents that US immigrants from most nationalities are positively selected on education). In terms of differential selection, if anything the relationship is negative, especially when weighted by the number of second generation immigrants in the sample (the unweighted relationship is flatter, though still negatively sloped). Emigrant mothers from high PISA countries are more adversely selected in terms of education than those from low PISA countries (panel 3 of Figure 1.3).[21]

[19] Ideally, we would like to perform such an exercise with a measure of quality pre-determined with respect to migration. Parental education, as any other socio-economic control available in the PISA dataset, does not satisfy this condition, since parents might have acquired part of their education in their host countries, or have based their educational choices in their countries of origin anticipating their future relocation. In the Census data it is possible to alleviate these concerns by focusing on parents that completed their education in the country of origin (see Appendix A.3). However, this "contamination" of our proxy for unobservable parental skills is problematic for our purposes only to the extent that is differential across countries of origin.

[20] We construct a mapping between the ISCED classification of educational levels and equivalent years of schooling by using the country-specific conversion table in OECD (2012b).

[21] The results for China might appear in contrast with Feliciano (2005a), which argues that Chinese immigrants in the US are among the most positively selected in terms of education. Indeed, in Appendix A.3 we show that in the US Census data, while on average the pattern of differential selection with respect to the PISA score is still negative, the Chinese are relatively positively selected. This discrepancy is explained by the fact that the PISA test is only administered in Shanghai, and as such it targets a subsample of the Chinese population significantly more educated than the average. While Chinese emigrants are positively selected compared to Chinese non-emigrants

In addition, Table 1.4 shows results of a regression of our measure of selection of emigrant parents on the average PISA score in their country of origin, controlling for country (columns 1 and 3) and school (columns 2 and 4) fixed effects. For both mothers and fathers, the point estimates are negative and not statistically significant, suggesting that the type of differential selection that would invalidate our results is not present neither within host countries nor within schools.

We conclude this section by noticing that our results are consistent with the patterns of selection reported in the development accounting literature. Schoellman (2012) documents that, among migrants (not necessarily parents of school-age children) residing in the US, the education gap compared to non-migrants is higher for poor origin countries. Hendricks and Schoellman (2016) show that emigrants from poor countries are more positively selected in terms of pre-migration wages and occupations.[22]

1.6 Decomposition

This section quantifies the role of parental influence in accounting for cross-country differences in average test scores. For this purpose, we introduce a more general model, which allows both maternal and paternal influence to differ across countries and includes both natives and second-generation immigrants.[23] Suppose that the test score in wave t of student i, educated in school s and country c, whose mother and father were born in countries m and f is given by

$$T_{icst}^{mf} = Parents_{icst}^{mf} + \theta^m NatMoth_{icst}^{mf} + \zeta^f NatFath_{icst}^{mf} + \rho' Z_{icst}^{mf} + \alpha_{cs} + \alpha_t + \varepsilon_{icst}^{mf} \quad (1.2)$$

where $Parents_{icst}^{mf}$ represents parental influence, and is given by

$$Parents_{icst}^{mf} = \gamma^m + \delta^f + \beta' ParentsEdu_{icst}^{mf} + \lambda' ParentsOcc_{icst}^{mf} + \eta_{icst}^{mf} \quad (1.3)$$

with γ^m and δ^f being country-specific components capturing a set of average (unobservable) characteristics of mothers and fathers from countries m and f respectively. Parental influence includes also the effect of parents' education and occupational status.[24]

In (1.2), $NatMoth_{ist}^{mf}$ and $NatFath_{ist}^{mf}$ are dummies identifying native parents (mothers and fathers, respectively). The coefficient θ^m (and similarly ζ^f), in the spirit of a difference in differences, captures the extent to which the relative performance of students whose mother is from country m, compared to second-generation immigrant students from another country,

as a whole, they are negatively selected compared to the population involved in the PISA test.

[22]More broadly, our results are consistent with a large literature studying the determinants of emigrants' self-selection, such as income inequality, migration costs, social networks, geography and school quality. See Appendix A.3 for a detailed discussion.

[23]While the linear specification considered here is still restrictive, in Appendix A.5 we show that considering dfferent types of complementarities does not majorly alter our conclusions.

[24]Specifically, $ParentsEdu_{icst}^{mf}$ contains dummies for primary, secondary or tertiary education for each parent, while $ParentsOcc_{icst}^{mf}$ contains dummies for employment status (full-time employed, part-time employed, not working), as well as interactions between the full-time employed and part-time employed dummies and the ISEI index of occupational status.

is larger or smaller in country m (where the mother is native) as opposed to a different host country. We allow the "native advantage" to be country-specific for both mothers and fathers: a failure to do so would imply that this kind of variation would be absorbed by the country of origin fixed effects (see footnote 28 for further discussion on this point). In addition, Z_{icst}^{mf} includes controls for student's gender and age in months.

Differences in school quality are captured by either host country or school fixed effects.[25] The distinction between the two specifications is important, given that, within the same host country, students from high PISA countries seem to attend better schools. This within-country variation in school quality is attributed to parental influence in the host country fixed effect specification, but not when school fixed effects are introduced. While the selection of better schools is one of the channels through which parental influence manifests itself, for the purpose of explaining differences in the average performance of natives across countries, the extent to which differences in the average ability or willingness to select better schools can matter is limited by the available supply of school quality in each country.[26] We display results from both specifications, with the understanding that controlling for school fixed effects provides us with a lower bound for the importance of parental influence.

Combining (1.2) and (1.3) we obtain our main specification

$$\begin{aligned} T_{icst}^{mf} &= \beta' ParentsEdu_{icst}^{mf} + \lambda' ParentsOcc_{icst}^{mf} + \gamma^m + \delta^f \\ &+ \theta^m NatMoth_{icst}^{mf} + \zeta^f NatFath_{icst}^{mf} + \rho' Z_{icst}^{mf} + \alpha_{cs} + \alpha_t + u_{icst}^{mf} \end{aligned} \tag{1.4}$$

The object whose variation we are ultimately interested in decomposing is the average score (across all available waves) of native students in country c, which is

$$T^c = \alpha + Parents^c + \theta^c + \zeta^c + \overline{\alpha}_c + \rho' \overline{Z}_c \tag{1.5}$$

where $Parents^c = \gamma^c + \delta^c + \beta' \overline{ParentsEdu}^c + \lambda' \overline{ParentsOcc}^c$, $\overline{\alpha}_c$ is either a weighted average of the school fixed effects or the fixed effect for host country c (depending on the specification) and $\overline{Z}_c$, $\overline{ParentsEdu}^c$ and $\overline{ParentsOcc}^c$ are within-country c averages.[27] Equation (1.5) makes our decomposition explicit: our objective is to evaluate the importance of $Parents^c$ to account for the variation of T^c across countries.[28]

[25]These fixed effects also absorb the impact of any other institutional factor that influences directly or indirectly students' performance. Importantly, this includes all parental influences driven by context-specific factors, which, as discussed in section 1.2, are not part of our parental component identified out of second-generation immigrants.

[26]At one extreme, if all schooling resources are utilized to full capacity, endowing a country with a higher average parental effectiveness in school selection would not contribute at all to boosting the average score. This scenario is probably too stark, since in several countries students might be able to access better schools without necessarily displacing others, or parents' drive for school quality could stimulate its supply to start with.

[27]The constant α absorbs a weighted average of the wave fixed effects.

[28]Notice that θ^c and ζ^c are not included in $Parents^c$. These parameters are identified out of the comparison between native and second-generation immigrant students in country c, and we think that various factors different from parental influence (such as the extent to which immigrants manage or are willing to integrate in their host country, or even characteristics of the school curriculum) could drive the international variation in the "native advantage". Instead, we view our focus on second-generation immigrants of different nationalities as one of the main advantages of our empirical approach, as it enables us to clean our estimates from confounders that would be difficult to proxy for. Nevertheless, θ^c and ζ^c are both positively correlated with T^c, so including them in our

We estimate the country-c-specific parental component from

$$\widehat{Parents}^c = \widehat{\gamma}^c + \widehat{\delta}^c + \widehat{\beta}'\overline{ParentsEdu}^c + \widehat{\lambda}'\overline{ParentsOcc}^c$$

where $\widehat{\gamma}^c, \widehat{\delta}^c$, $\widehat{\beta}$ and $\widehat{\lambda}$ are estimated from (1.4). As discussed, we focus on two specifications, one that includes school fixed effects and another with host country fixed effects only.

Figure 1.5 plots the parental component obtained from both specifications against the average score of natives ($Parents^{\text{CHINA}}$ is normalized to 1 in both cases).[29] $Parents^c$ is larger (in absolute terms) for high scoring countries, which means that parental influence does account for some of the cross-country variation (as opposed to masking an even larger dispersion) in performance. The dispersion in $Parents^c$ is larger for the country fixed effect specification, which allows the parental component to absorb the within-country variation in school quality.

As a simple summary statistic, we define the share of the total cross-country variance of T^c accounted by $Parents^c$ as[30]

$$V_{Parents} = \frac{Var(Parents^c)}{Var(T^c)} \tag{1.6}$$

This can be interpreted as the fraction of the variance that would persist if all relevant factors except parental influence were equalized across countries. To evaluate the relative contribution of observable and unobservable parental characteristics, we also compute an equivalent statistic for the country-specific intercepts only,

$$V_{FE} = \frac{Var(\gamma^c + \delta^c)}{Var(T^c)} \tag{1.7}$$

As a result of sampling error, the variance of our estimates overstates the true variation in the corresponding quantities. This is particularly relevant for $Parents^c$ and $\gamma^c + \delta^c$, which for some countries are identified out of a limited number of second generation immigrants. As suggested by Aaronson et al. (2007), we adjust variances by subtracting the average squared standard error of our estimates.[31]

Table 1.5 shows the results of this variance decomposition. According to the estimates adjusted for sampling error, $Parents^c$ accounts for at least 14% of the cross-country variance, and up to 20% when we do not clean it from the variation in school quality within countries. Most of the variation in $Parents^c$ is driven by $\gamma^c + \delta^c$, suggesting that cross-country differences in parents' education and occupational status are of limited quantitative importance. The adjustment for sampling error approximately halves the inferred contribution of $Parents^c$ and $\gamma^c + \delta^c$

parental component would lead us to infer a (moderately) higher role for parental influence.

[29]Table 1.6 displays $Parents^c$ for all countries.

[30]Our decomposition exercise is similar to the ones proposed in Card et al. (2013) and Finkelstein et al. (2016), who also use (in different contexts) fixed effects identified out of movers to separate the contribution of individual characteristics and geographical or institutional factors.

[31]Standard errors are computed using the provided replicate weights, and inflated to account for the estimated measurement error in test scores. For computational convenience, we used the "unbiased shortcut" procedure described in OECD (2009). See Appendix A.2 for more details on the construction of standard errors with PISA data.

compared to the unadjusted estimates (shown in the first row of Table 1.5).

We then investigate the contribution of parental influence for the out-performance of Chinese students. For each country c we define the share of the gap in average test score with respect to China accounted by parental influence as

$$S_{Parents}(c, \text{CHINA}) = \frac{Parents^{\text{CHINA}} - Parents^{c}}{T^{\text{CHINA}} - T^{c}} \tag{1.8}$$

Moreover, as in (1.7), we isolate the importance of unobservable parental characteristics by computing

$$S_{FE}(c, \text{CHINA}) = \frac{(\gamma^{\text{CHINA}} + \delta^{\text{CHINA}}) - (\gamma^{c} + \delta^{c})}{T^{\text{CHINA}} - T^{c}} \tag{1.9}$$

Table 1.6 shows that parental influence plays a substantial role in accounting for the gap between China and the rest of the world. On average, between 22% and 58% of China's out-performance can be accounted for by parental influence. The gap between the school and country fixed effect specifications suggests that school choice is a particularly important factor that sets Chinese parents apart. Virtually all the gap in parental influence is driven by the country-specific intercepts. While some of the country-specific estimates are too imprecise to allow definite conclusions, the gaps in $Parents^{c}$ are particularly high for several countries in the middle-bottom part of the score distribution, but not so pronounced for the worst performers.

The results are particularly striking for Southern European countries, which display large gaps with respect to China in terms of both test scores and parental influence. On the other hand, parental influence plays a limited role for Latin American countries, whose poor performance in standardized test has been object of recent study (Hanushek and Woessmann, 2012b).[32]

1.7 Mechanism

We now study the nature of differences in parental influence. What is it about parents from high PISA countries that drives the superior school performance of their children? While answering this question precisely is difficult, we make progress by proceeding in three steps. First, we distinguish between several possible drivers of cross-country differences in parental influence: the educational system to which parents were exposed, the country-specific cultural context and the genetic transmission of relevant traits. Then we turn to time use surveys to see whether immigrant parents from high PISA countries differ in observable practices that might help to explain their children's better performance at school. Finally, we test whether country level proxies for economic development, educational attainment or culture can explain our relationship of interest.

[32] In Appendix A.6 we show in a standard development accounting framework how the relative variation in parental influence maps into its relative contribution for cross-country differences in output. We find that the parental component accounts for about 12% of the covariance between GDP per worker and the PISA score.

1.7.1 Interactions

Cross-country differences in parental influence might be driven by a number of sources. First, the outstanding performance of second-generation immigrants from high PISA countries might reflect the higher quality of the education received by parents in their country of origin. While this would still imply that these students have an advantage in terms of parental influence, the source of this advantage would be the school system in the parents' country of origin, implying a powerful intergenerational multiplier effect of educational quality. This would provide an even stronger rationale for policies aiming to replicate the best practices in this domain.

An alternative is that the country-of-origin's cultural context, defined as a shared set of beliefs and preferences within a given country, might have shaped parents' attitudes and beliefs towards education. This variation in cultural traits might have its roots in factors deeply entrenched in a country's history and culture, and improving the educational system might not do much in raising average test scores if these aspects do not change as well. Yet another possibility is that individuals from different countries are systematically endowed with different genetic traits that shape their human capital investment. This interpretation would leave little room for policies to affect achievement gaps.

To discriminate between these views, we explore the heterogeneity of country-specific parental influences with respect to parental characteristics. If the intergenerational transmission of educational quality is important, we expect the correlation between school performance and the PISA score in the parents' country of origin to be particularly strong for students whose parents acquired more education in their home country, and were therefore more exposed to the educational system.[33] At the extreme, parents with no education cannot transmit the quality of their home country's school system at all.

On the other hand, if what matters is the cultural context in the source country, we expect the country of origin effect to be smaller among parents that are more integrated in their host country and have at least in part converged to its cultural norms. As cultural assimilation takes time, the correlation between children's performance and the average test score in the country of origin should be weaker for parents that have emigrated many years ago.[34] Moreover, there is evidence that highly educated immigrants have an easier time integrating in their host country (Lichter and Qian, 2001; Meng and Gregory, 2005); therefore, under the "cultural" interpretation parental years of schooling should also alleviate the correlation between children's performance and the average score in parents' country of origin.

To summarize, we have testable implications to discriminate between two sources of dif-

[33] We might expect a differential effect of years of schooling in the host country as well, if there are dynamic complementarities in the human capital accumulation process that make the impact of an additional year of schooling stronger for parents that have spent the initial part of their educational career in higher quality schools. Moreover, emigrants from high PISA countries might attend better schools once in the host country.

[34] There is widespread evidence that years since migration correlate positively with immigrants' assimilation (Chiswick, 1978). Children of parents that have spent more time in the US fare better in terms of years of schooling, earnings (Abramitzky et al., 2016) and school performance (Nielsen and Schindler Rangvid, 2011), a result that we confirm in our setting (with the caveat that the impact of years since migration is heterogeneous depending on the country of origin). Appendix A.4 shows that results are similar when we focus on alternative measures on immigrants' assimilation.

ferences in parental influence. The intergenerational transmission of educational quality mechanism would imply a positive interaction term between parental years of schooling acquired in the home country and the average score of natives in the same country. A story based on differences in cultural environments would instead involve a negative interaction between the average test score and parents' years since migration, as well as parents' years of schooling. A purely genetic view, instead, would not have any obvious implication in terms of differential effects.

We now turn to the US Census data to put these predictions to empirical scrutiny. We compute mothers' years of schooling both in their home and in their host countries based on information on year of immigration and age at the end of education (imputed from the educational level).[35]

Table 1.7 shows the results. We add to the baseline specification in column 1 an interaction term between T^m and mother's years of schooling, finding a negative and significant coefficient (column 2). When we break down years of schooling between those acquired in the US and those acquired in country m (column 3), we find that the interaction term is negative in both cases, with coefficients of similar magnitudes. Figure 1.6 plots the coefficient on T^m for different levels of mothers' educational attainment: most of the gap is driven by mothers with either no education or primary schooling only, and disappears when we focus on mothers with college education. These results are inconsistent with strong intergenerational effects of educational quality.[36]

The study of the heterogeneity with respect to years since migration supports the importance of country-specific cultural environments. According to column 4 in Table 1.7, the correlation between T^m and children's school performance is weaker for mothers that have emigrated many years ago.[37] As shown in Figure 1.7, the effect of T^m disappears for mothers that have spent 25 years in the US, suggesting that a relatively quick convergence of cultural norms might be taking place. Column 5 shows that this pattern (as well as the results on education discussed above) is unaffected by the inclusion of controls for age at migration, which has also been shown to be important for the assimilation of immigrants (Bleakley and Chin, 2010).

A possible concern is that the imperfect mapping from the information available in the Census to years of schooling accumulated in country m and in the US might confound our results. Column 6 in Table 1.7 shows results for a sub-sample of mothers entirely educated in

[35]Year of immigration is available as a categorical variable, in intervals of approximately 5 years. We impute the exact year of arrival in the US according to two alternative criteria: the middle year of each interval for our baseline results, and the first year for a robustness check where we consider parents likely to have completed their education in their origin country.

[36]It is interesting to contrast these results to the ones in Schoellman (2012), who shows that the wage returns to education of US immigrants are positively related to GDP per capita and PISA scores in their home country and interprets this as evidence in favor of the fact that school quality varies across countries. While differences in school quality might be important for immigrants' labor market outcomes, they do not seem to account for the differential school performance of their children.

[37]This result provides an additional reason why our decomposition exercise in Section 1.6 might understate the importance of parental influence. If immigrant parents from different countries progressively become more similar to each other as they integrate in their host country, we would find a larger role for parental influence by focusing on those who have just emigrated, which are still very comparable to non-emigrants in their country of origin. Unfortunately, date of immigration is not available in the PISA data.

their country of origin. The interaction between T^m and mother's years of schooling is negative and significant, and so is the one between T^m and years since migration. The magnitudes of the estimated coefficients are virtually identical to the ones obtained with the full sample.

Overall, our results are supportive of an interpretation based on country-specific cultural environments. While we cannot entirely rule out a role for genetic traits, the fact that gaps in performance disappear when focusing on more educated and integrated parents is difficult rationalize with a purely genetic transmission story.

1.7.2 Time Use

In this section we investigate whether immigrant parents from high PISA countries allocate more time to activities that might plausibly stimulate their children's human capital accumulation. The analysis complements and extends the work of Ramey (2011), who compares time use practices across ethnic groups.

Table 1.8 shows our results. Columns 1 to 3 refer to total child care, while columns 4 to 6 break down the time spent with children in the educational, recreative and basic categories. Across all specifications and time use categories, interviewed parents from high PISA countries stand out for spending more time with their children. The result is robust to the inclusion of state fixed effects and several controls on demographic and socio-economic characteristics of both parents and children. Since time use variables are measured in minutes and refer to a single day, from column 3 it emerges that an increase of one (individual-level) standard deviation in the PISA score in a parent's country of origin corresponds to a higher investment of approximately 57 minutes per week in total child care. This extra child care time is quite evenly spread across the three time use subcategories, even though as a proportion of the mean the largest gap is in educational activities. [38]

These results indicate that immigrant parents do differ in terms of observable practices as a function of their country of origin and this may lay behind the results found in the previous sections.

1.7.3 Country-Level Characteristics

We now augment specification (1.1) with a series of controls at the mother's country of origin-level. The objective is to verify whether the estimate of our coefficient of interest picks up variation across country-level characteristics that might plausibly affect second-generation immigrants' school performance.

Table 1.9 includes controls related to economic development and schooling in country m. As high-scoring countries in the PISA test are richer and have more educated populations, we want to check whether this gives to second-generation immigrants from those countries some direct advantage which might explain their superior performance. In columns 2 and 3 we add

[38] We find again that the effect is mostly concentrated among low educated parents and those of more recent immigration. These results are available upon request.

to the baseline specification in column 1 controls for contemporaneous log real GDP per capita and average years of schooling in the mother's country of origin; the coefficients are small and not statistically significant. In column 5 we further control for the log expenditure per pupil in secondary schools; once again, compared to the baseline regression on the same sample reported in column 4, the added regressor has negligible explanatory power and our coefficient of interest is not affected.[39]

Table 1.10 controls for proxies for various cultural traits from the World Value Survey. While, to our knowledge, a direct measure of attitudes towards education is not available, we focus on three proxies that have been studied elsewhere as determinants of labor supply and effort: tastes for leisure, locus of control and long-term orientation.[40,41]

Columns 1 to 3 introduce our cultural proxies in regressions controlling for the usual parental characteristics and school fixed effects. All three coefficients are significant and of the expected sign; second-generation immigrants from countries where leisure is considered less important, where people believe to have control on events in their life and are oriented towards the future score better than their peers, even if school quality is controlled for. A similar message emerges when the cultural proxies are included simultaneously (column 4). In column 6 we further control for the average performance of native students in the mother's country of origin, which retains its statistical significance and drops by one third compared to the baseline specification without cultural proxies (reported in column 5).

Proxies for cultural traits in the parental countries of origin can go some way towards explaining the parental country of origin effect across second-generation immigrants. Much of this variation, however, remains unexplained, suggesting that the attitudes or traits underlying educational performance might not entirely captured by the proxies for culture commonly used in the literature.

1.8 Conclusions

While the quality of the educational system and local economic conditions are often named as the key factors for cross-country differences in human capital, this is not the whole story. We show that an important share of the international variation in test scores is driven by cross-country differences in broadly defined parental influence. We arrive to this conclusion through

[39]We use average years of schooling for 35- to 45-year-old adults in 2005, and expenditure per pupil in secondary schools in 2000 (the year with the largest number of observations in Bartik (2008)'s dataset). Using different years and reference groups yields very similar results.

[40]Among others, Moriconi and Peri (2015) study country-specific preferences for leisure and labor supply choices, Coleman and DeLeire (2003) estimate the effect of the locus of control on educational and labor market outcomes while Dohmen et al. (2016), Galor and Ömer Özak (2016) and Figlio et al. (2016) consider how long term orientation shapes human capital investment.

[41]Tastes for leisure are measured from the question *how important leisure time is in your life?*. Answers (ranging from 1 to 4) are standardized to take mean 0 and standard deviation 1 at the individual level. The locus of control is measured from the question *how much freedom of choice and control you feel you have over the way your life turns out?*, where answers are standardized as above. The measure of long term orientation was developed by Hofstede (1991) and updated in Hofstede et al. (2010) using data from the World Value Survey; it ranges from 0 to 1.

an indirect empirical approach, based on the comparison between the performance of second-generation immigrants with parents of different nationalities. Parental influence operates both within schools and through school choice, highlighting potentially important interactions between parental and schooling inputs for human capital formation.

We do not find evidence for a mechanism of intergenerational transmission of school quality, as parental education appears to attenuate rather than reinforce the relevance of the standardized test performance in parents' country of origin for explaining their children's achievements. Our results support instead the importance of cultural factors, varying across countries, that shape parents' attitudes towards their children's education. Differences in parental influence across nationalities are partially reflected in observable time use practices.

These results have important implications for the study of human capital in a cross-country perspective. Models of human capital accumulation should be consistent with an important role for parents in the transmission of knowledge. Moreover, parental attitudes towards education potentially represent a competing mechanism to gaps in TFP and local economic conditions for generating human capital and output gaps across countries. A systematic quantitative analysis of the interaction between these factors is left for future work.

Our paper opens other important avenues for future research. If parental attitudes towards education are important determinants of human capital achievement, it is crucial to understand how they form and evolve, and why they do so differently across time and space. Historical circumstances experienced in different countries might have played an important role, and social interactions between people of various origins (brought about by migration or trade linkages) might have shaped the diffusion of different cultural traits. Further research is also needed to identify the specific activities, attributes or skills responsible for the cross-country variation in parental influence.

Our results are relevant for policymakers aiming to raise their students' performance in standardized tests. Cross-country gaps go beyond differences in school quality, and policies aimed at replicating school practices successful in other countries might be uneffective. Parents are an important factor, and it is an open question whether policy should and could play a role in this respect.

1.9 Tables and Figures

Table 1.1: Summary statistics - Second Generation Immigrants on the Mother's Side

Panel A: PISA Sample	All		Score Country *m* Below Median		Score Country *m* Above Median	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Score	0.29	0.92	0.10	0.87	0.75	0.88
Score Country *m*	0.28	0.56	-0.01	0.29	0.98	0.39
Mother Sec Edu	0.51	0.50	0.50	0.50	0.56	0.50
Mother Ter Edu	0.31	0.46	0.35	0.48	0.22	0.41
Father Sec Edu	0.51	0.50	0.49	0.50	0.55	0.50
Father Ter Edu	0.34	0.47	0.38	0.49	0.25	0.43
Working Mother ISEI	41.35	18.79	41.33	19.08	41.40	17.99
Working Father ISEI	41.51	17.42	41.32	17.40	41.97	17.46
Immigrant Father	0.64	0.48	0.66	0.47	0.59	0.49
Observations	40067		20320		19747	

Panel B: US Census Sample	All		Score Country *m* Below Median		Score Country *m* Above Median	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
No Grade Repeated	0.81	0.39	0.76	0.43	0.85	0.36
Score Country *m*	0.48	0.50	0.04	0.35	0.87	0.22
Mother Sec Edu	0.48	0.50	0.34	0.47	0.61	0.49
Mother Ter Edu	0.21	0.40	0.14	0.35	0.26	0.44
Father Sec Edu	0.39	0.49	0.32	0.47	0.46	0.50
Father Ter Edu	0.34	0.47	0.23	0.42	0.43	0.49
Log Family Income	10.84	0.69	10.68	0.73	10.98	0.62
Immigrant Father	0.46	0.50	0.63	0.48	0.31	0.46
Yrs Since Migr Mother	20.08	8.75	19.20	8.85	20.84	8.59
Student Age	11.35	2.29	11.21	2.29	11.46	2.28
Observations	53081		27071		26010	

Notes: The Table shows descriptive statistics for second generation immigrants on the mother's side in the PISA (Panel A) and US Census (Panel B) samples. Only cases where both parents report a country of origin and the country of origin of the mother participates to PISA are included. Scores are from the math test and are standardized to have mean 0 and (individual-level) standard deviation 1 across the (pooled, equally weighted) 73 countries participating to the test. Observations weighted according to the provided sample weights.

Table 1.2: Main results - PISA

	Dependent Variable: Math Test Score					
	[1]	[2]	[3]	[4]	[5]	[6]
			All			No East Asia
Score Country *m*	0.662***	0.499***	0.253***	0.249***	0.240***	0.225**
	(0.081)	(0.158)	(0.073)	(0.070)	(0.065)	(0.095)
Female	-0.140***	-0.148***	-0.206***	-0.204***	-0.201***	-0.187***
	(0.032)	(0.028)	(0.022)	(0.022)	(0.022)	(0.024)
Father Sec Edu				0.030	0.014	0.022
				(0.022)	(0.022)	(0.044)
Father Ter Edu				0.099***	0.045	0.049
				(0.033)	(0.034)	(0.052)
Mother Sec Edu				0.001	-0.015	0.027
				(0.037)	(0.037)	(0.065)
Mother Ter Edu				0.032	-0.011	0.023
				(0.042)	(0.042)	(0.075)
Mother Working $\times$ Mother ISEI					0.003***	0.003***
					(0.001)	(0.001)
Father Working $\times$ Father ISEI					0.003***	0.003***
					(0.001)	(0.001)
N	40067	40067	40067	40067	40067	25454
# Country m	49	49	49	49	49	42
R Squared	0.16	0.25	0.67	0.67	0.67	0.63
Host Country FE	No	Yes	Yes	Yes	Yes	Yes
School FE	No	No	Yes	Yes	Yes	Yes

Notes: The Table shows results for second generation immigrants on the mother's side. The sample includes only cases where both parents report a country of origin and the country of origin of the mother participates to PISA. *Score Country m* is the average math PISA score of natives (standardized to have mean 0 and standard deviation 1 across all countries participating to the test) in the country of birth of the mother, across all available waves. All specifications control for intercept, students' age (in months), wave fixed effect and a dummy for father's immigrant status; specifications 5-6 additionally control for dummies for parents' employment status (full-time employed, part-time employed, not working). *Working* refers to either full-time or part-time employed. Observations are weighted according to the provided sample weights. Standard errors are clustered by mother's country of origin, and inflated by the estimated measurement error in test scores. * denotes significance at 10%, ** at 5%, *** at 1%.

Table 1.3: Main results - US CENSUS

	Dependent variable: 1 = Never repeated a grade					
	[1]	[2]	[3]	[4]	[5]	[6]
			All			No East Asia
Score Country *m*	0.088***	0.059***	0.034***	0.031***	0.028***	0.022*
	(0.030)	(0.017)	(0.009)	(0.010)	(0.009)	(0.012)
Female	0.068***	0.067***	0.067***	0.068***	0.068***	0.071***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Mother Sec Edu			0.054***	0.051***	0.047***	0.045***
			(0.013)	(0.012)	(0.011)	(0.012)
Mother Ter Edu			0.068***	0.064***	0.054***	0.050***
			(0.010)	(0.010)	(0.010)	(0.010)
Father Sec Edu			0.041***	0.041***	0.036***	0.041***
			(0.012)	(0.011)	(0.010)	(0.009)
Father Ter Edu			0.072***	0.073***	0.058***	0.063***
			(0.015)	(0.014)	(0.011)	(0.011)
Log Family Income					0.036***	0.037***
					(0.008)	(0.009)
N	53081	53081	53081	53081	53081	49132
# Country *m*	61	61	61	61	61	54
R Squared	0.06	0.09	0.10	0.10	0.10	0.11
Comm Zone FE	No	Yes	Yes	Yes	Yes	Yes
Years Since Migr Mother	No	No	No	Yes	Yes	Yes

Notes: The Table shows results for second generation immigrants on the mother's side. *Score Country m* is the average math PISA score of natives (standardized to have mean 0 and standard deviation 1 across all countries participating to the test) in the country of birth of the mother. All specifications control for intercept, child age dummies, parents' age, number of siblings, year fixed effect, (year specific) quarter of birth fixed effect and father's immigrant status. Observations weighted according to the provided sample weights. Robust standard errors clustered by mother's country of origin. * denotes significance at 10%, ** at 5%, *** at 1%.

Table 1.4: Selection

	Dependent Variable: Standardized Years of Education			
	[1]	[2]	[3]	[4]
	Mothers		Fathers	
Score Country *m*	-0.072	-0.219		
	(0.193)	(0.147)		
Score Country *f*			-0.093	-0.208
			(0.193)	(0.145)
N	40067	15710	40304	40304
R Squared	0.12	0.64	0.13	0.59
Host Country FE	Yes	Yes	Yes	Yes
School FE	No	Yes	No	Yes

Notes: The sample includes emigrant mothers (columns 1 and 2) and fathers (3 and 4). The dependent variable is years of education standardized by the average and standard deviation of mothers' (columns 1 and 2) and fathers' (3 and 4) education in the country of origin. *Score Country m* and *Score Country f* are the average math PISA scores of natives (standardized to have mean 0 and standard deviation 1 across all countries participating to the test) in the country of birth of the mother and the father. All specifications control for intercept and wave fixed effect. Standard errors clustered by mother's (columns 1 and 2) and father's (3 and 4) country of origin. * denotes significance at 10%, ** at 5%, *** at 1%.

Table 1.5: Decomposition Results - Cross-Country Variance

	$V_{Parents}$ (%)		V_{FE} (%)	
	School FE	Host Country FE	School FE	Host Country FE
Unadjusted	24.25	34.65	23.22	30.59
Adjusted	14.13	19.94	13.10	15.91

Notes: The Table shows the ratio (in percent) between the cross-country variance of either the whole parental component ($V_{Parents}$) or the country specific intercept (V_{FE}) and the cross-country variance of the average math PISA score of natives. Columns denoted by *School FE* (*Host Country FE*) refer to specifications that include school fixed effects (host country fixed effects). Adjusted variances are computed by subtracting the average squared standard errors (constructed using the provided replicate weights, and inflated by the estimated measurement error in test scores).

Table 1.6: Decomposition Results - Countries

Country	PISA Score	$Parents^c$		$S_{Parents}(c, \text{CHINA})$ (%)		$S_{FE}(c, \text{CHINA})$ (%)	
		School FE	Country FE	School FE	Country FE	School FE	Country FE
China	1.33	1	1	-	-	-	-
Hong Kong	0.92	0.91	0.79	22.39	52.15	20.22	41.11
				(25.66)	(31.83)	(25.69)	(31.76)
Switzerland	0.82	0.95	0.56	10.42	86.15	15.33	101.30
				(20.63)	(30.10)	(20.65)	(30.29)
Belgium	0.79	0.91	0.43	16.16	104.93	22.02	122.00
				(27.44)	(35.82)	(27.37)	(36.22)
Netherlands	0.74	0.92	0.75	12.71	42.65	18.72	59.06
				(14.76)	(20.78)	(14.94)	(21.03)
Germany	0.64	0.90	0.62	14.08	54.13	17.47	64.04
				(9.20)	(15.54)	(9.22)	(15.88)
New Zealand	0.58	0.66	0.29	45.30	95.09	48.11	102.03
				(8.84)	(11.36)	(8.97)	(11.84)
Estonia	0.55	0.93	0.70	8.90	38.38	10.50	43.78
				(25.76)	(29.80)	(25.79)	(29.82)
Macao	0.55	0.95	0.74	6.68	33.07	4.32	21.98
				(11.51)	(14.04)	(11.49)	(14.04)
France	0.52	0.77	0.35	27.80	79.93	29.23	84.99
				(7.17)	(10.21)	(7.26)	(10.50)
Australia	0.50	0.63	0.44	45.01	67.68	48.45	76.76
				(16.63)	(16.39)	(16.62)	(16.56)
Denmark	0.50	0.99	0.61	1.01	46.57	4.13	56.13
				(20.69)	(20.39)	(20.65)	(20.42)
Austria	0.48	0.79	0.26	24.10	87.41	25.15	91.85
				(14.01)	(18.45)	(14.08)	(18.75)
Czech Republic	0.46	0.73	0.37	31.17	72.36	31.06	72.64
				(13.32)	(19.85)	(13.32)	(19.84)
Sweden	0.44	0.85	0.48	16.60	57.55	20.17	68.25
				(9.03)	(10.89)	(9.02)	(11.06)
Vietnam	0.44	0.87	0.49	14.92	57.18	2.79	23.78
				(8.74)	(9.43)	(8.66)	(9.53)
United Kingdom	0.42	0.73	0.38	30.11	67.52	32.09	73.71
				(5.27)	(7.35)	(5.34)	(7.56)
Poland	0.34	0.67	0.36	32.98	64.81	29.61	57.33
				(7.15)	(10.10)	(7.15)	(10.12)
Slovakia	0.33	0.80	0.35	19.65	65.13	17.87	61.59
				(9.48)	(11.97)	(9.51)	(12.15)
United States	0.26	0.96	0.79	4.04	20.00	7.59	29.26
				(9.46)	(10.28)	(9.48)	(10.43)
Spain	0.25	0.62	0.23	35.05	71.03	32.58	63.92
				(6.85)	(9.34)	(6.86)	(9.42)
Portugal	0.16	0.65	0.22	29.89	67.05	25.14	51.60
				(4.99)	(7.26)	(5.00)	(7.32)
Italy	0.14	0.54	-0.01	38.35	84.36	37.27	82.02
				(5.98)	(7.70)	(6.02)	(7.77)
Russia	0.12	0.80	0.54	16.85	38.11	17.77	41.79
				(4.11)	(6.19)	(4.13)	(6.23)
Croatia	0.06	0.57	0.22	33.40	61.57	32.60	60.59
				(8.98)	(11.64)	(8.97)	(11.68)
Greece	-0.02	0.46	0.01	39.88	73.52	39.65	72.61
				(10.17)	(11.39)	(10.18)	(11.44)
Turkey	-0.21	0.39	-0.34	39.35	87.29	34.42	71.19
				(3.78)	(5.66)	(3.79)	(5.29)
Serbia-Mont.	-0.23	0.55	0.05	28.87	60.64	28.61	60.87

Continued on next page

Continued from previous page

Country	PISA Score	$Parents^c$		$S_{Parents}(c, \text{CHINA})$ (%)		$S_{FE}(c, \text{CHINA})$ (%)	
		School FE	Country FE	School FE	Country FE	School FE	Country FE
				(3.17)	(4.97)	(3.14)	(4.96)
Romania	-0.31	0.62	0.35	23.16	39.56	21.56	36.88
				(7.08)	(8.33)	(7.10)	(8.32)
Uruguay	-0.34	0.87	0.40	7.50	36.03	5.09	29.26
				(12.17)	(17.40)	(12.15)	(17.45)
Chile	-0.38	0.65	0.28	20.40	42.27	17.35	35.39
				(10.58)	(15.27)	(10.59)	(15.30)
Malaysia	-0.41	0.67	0.01	18.63	56.42	16.71	51.66
				(11.37)	(18.10)	(11.43)	(18.24)
Argentina	-0.63	0.85	0.47	7.45	26.78	5.96	22.39
				(8.54)	(12.06)	(8.53)	(12.08)
Jordan	-0.67	0.59	0.14	20.23	43.11	19.62	41.68
				(3.90)	(4.98)	(4.04)	(5.22)
Albania	-0.68	0.45	-0.02	27.19	50.65	24.64	44.71
				(3.15)	(4.38)	(3.15)	(4.42)
Brazil	-0.75	0.81	0.37	9.16	30.16	6.31	22.01
				(7.44)	(8.54)	(7.43)	(8.58)
India	-0.98	0.66	0.23	14.83	33.04	11.13	22.82
				(2.68)	(3.43)	(2.67)	(3.40)
Average	0.18	0.75	0.38	22.06	58.17	21.70	57.30
				(4.42)	(6.35)	(4.45)	(6.46)

Notes: The Table shows the decomposition results across countries. Only countries with at least 100 immigrant parents in the sample are shown. $Parents^c$ is the estimated parental component, normalized such that $Parents^{\text{CHINA}} = 1$. Standard errors (in parentheses) are computed using the provided replicate weights, and inflated by the estimated measurement error in test scores.

Table 1.7: Interactions - US CENSUS

	Dependent variable: 1 = Never repeated a grade					
	[1]	[2]	[3] All	[4]	[5]	[6] Mothers Educated in m
Score Country *m*	0.030***	0.097***	0.097***	0.150***	0.168***	0.159***
	(0.008)	(0.024)	(0.025)	(0.033)	(0.038)	(0.038)
Female	0.067***	0.067***	0.067***	0.067***	0.067***	0.068***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.005)
Yrs Edu Father	0.005***	0.005***	0.005***	0.005***	0.005***	0.006***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Yrs Edu Mother	0.006***	0.007***				
	(0.001)	(0.001)				
Score Country *m* × Yrs Edu Mother		-0.006***				
		(0.002)				
Yrs Edu Mother in US			0.007***	0.003**	0.006***	
			(0.001)	(0.002)	(0.001)	
Yrs Edu Mother in *m*			0.007***	0.007***	0.007***	0.006***
			(0.001)	(0.001)	(0.001)	(0.001)
Score Country *m* × Yrs Edu Mother in US			-0.007***	-0.003*	-0.003**	
			(0.001)	(0.001)	(0.002)	
Score Country *m* × Yrs Edu Mother in *m*			-0.006***	-0.007***	-0.007***	-0.007***
			(0.002)	(0.002)	(0.002)	(0.002)
Yrs Since Migr Mother				0.003***	0.003***	0.004***
				(0.001)	(0.001)	(0.001)
Score Country *m* × Yrs Since Migr Mother				-0.003***	-0.003***	-0.003**
				(0.001)	(0.001)	(0.001)
Age Migration Mother					0.006***	0.005**
					(0.002)	(0.002)
Score Country *m* × Age Migration Moth					-0.001	-0.000
					(0.001)	(0.001)
N	53081	53081	53081	53081	53081	29963
# Country *m*	61	61	61	61	61	61
R Squared	0.10	0.10	0.10	0.11	0.11	0.12
Comm Zone FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The Table shows results for second generation immigrants on the mother's side. *Score Country m* is the average math PISA score of natives (standardized to have mean 0 and standard deviation 1 across all countries participating to the test) in the country of birth of the mother, across all available waves. All specifications control for intercept, child age dummies, parents' age, number of siblings, log family income, year fixed effect, (year-specific) quarter of birth fixed effect and father's immigrant status. Observations are weighted according to the provided sample weights. Standard errors are clustered by mother's country of origin. * denotes significance at 10%, ** at 5%, *** at 1%.

Table 1.8: Time Use of Parents

	Total	Total	Total	Educational	Recreational	Basic
	[1]	[2]	[3]	[4]	[5]	[6]
Score Country *p*	14.636*	12.822**	8.188**	2.208**	4.087**	1.894
	(8.489)	(6.319)	(3.448)	(1.100)	(1.711)	(1.857)
Mother			66.413***	8.449***	0.903	57.061***
			(4.000)	(0.885)	(3.179)	(2.416)
Parent Sec Edu			-2.355	4.482***	-2.827	-4.011*
			(5.617)	(0.674)	(3.285)	(2.138)
Parent Ter Edu			4.232	3.826***	-2.526	2.932
			(3.469)	(1.220)	(2.174)	(1.901)
Spouse Sec Edu			3.107	-1.783*	6.519**	-1.628
			(2.905)	(0.894)	(2.611)	(1.322)
Spouse Ter Edu			12.839***	2.409	7.242***	3.188
			(3.376)	(1.728)	(2.516)	(2.608)
Log Family Income			6.228***	0.719	-1.407	6.915***
			(2.140)	(0.630)	(0.959)	(1.353)
Age Parent			0.234	0.097	0.064	0.073
			(0.369)	(0.072)	(0.339)	(0.191)
Age Spouse			0.345	0.151	0.014	0.181
			(0.235)	(0.094)	(0.198)	(0.251)
Number of Children			20.072***	3.451**	1.003	15.617***
			(2.810)	(1.379)	(0.690)	(1.640)
Avg Age Children			-8.898***	-0.263*	-3.338***	-5.297***
			(1.065)	(0.141)	(0.439)	(0.577)
Number of Male Children			-1.138	0.849	-0.950	-1.036
			(1.680)	(0.545)	(1.046)	(1.031)
Yrs Since Migration			-0.162	-0.128***	-0.120	0.086
			(0.201)	(0.037)	(0.133)	(0.102)
N	5659	5659	5659	5659	5659	5659
# Country *p*	59	59	59	59	59	59
Mean Dep. Var.	89.87	89.87	89.87	10.53	22.27	57.07
St. Dev. Dep. Var.	119.98	119.98	119.98	32.30	58.06	88.63
R Squared	0.01	0.03	0.24	0.06	0.10	0.22
State FE	No	Yes	Yes	Yes	Yes	Yes
Year FE	No	Yes	Yes	Yes	Yes	Yes

Notes: The sample includes only immigrant parents of children of at most 18 years. $Parent$ refers to the interviewed parent, $Spouse$ to the other one; $Mother$ is 1 when the interviewed parent is the mother. $Total$ refers to the total time spent in child care activities, while $Educational$, $Recreational$ and $Basic$ refer to the sub-categories defined in the text. *Score Country* p is the average math PISA score of natives (standardized to have mean 0 and standard deviation 1 across all countries participating to the test) in the country of birth of the interviewed parent, across all available waves. Additional controls in specifications (3) to (6) are dummies for native spouses and for retired, full time students and disabled parents. Standard errors are clustered by the interviewed parent's country of origin. * denotes significance at 10%, ** at 5%, *** at 1%.

Table 1.9: Country of Origin Characteristics - Economic and Educational Development

	Dependent variable: Math Test Score				
	[1]	[2]	[3]	[4]	[5]
Score Country *m*	0.240***	0.254***	0.210***	0.214***	0.214***
	(0.065)	(0.061)	(0.063)	(0.065)	(0.063)
Female	-0.201***	-0.201***	-0.200***	-0.216***	-0.216***
	(0.022)	(0.022)	(0.022)	(0.019)	(0.019)
Father Sec Edu	0.014	0.012	0.013	0.014	0.014
	(0.022)	(0.022)	(0.022)	(0.023)	(0.023)
Father Ter Edu	0.046	0.044	0.045	0.053	0.053
	(0.034)	(0.034)	(0.034)	(0.041)	(0.041)
Mother Sec Edu	-0.015	-0.015	-0.022	-0.036	-0.036
	(0.036)	(0.037)	(0.036)	(0.032)	(0.032)
Mother Ter Edu	-0.012	-0.010	-0.020	-0.055	-0.055
	(0.042)	(0.042)	(0.045)	(0.041)	(0.042)
Mother Working $\times$ Mother ISEI	0.003***	0.003***	0.003***	0.003**	0.003**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Father Working $\times$ Father ISEI	0.003***	0.003***	0.003***	0.003**	0.003**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Log GDP Country *m*		-0.038			
		(0.036)			
Avg Years Edu in *m*			0.011		
			(0.011)		
Log Exp per Pupil in *m*					-0.003
					(0.024)
N	40029	40029	40029	31502	31502
# Country m	48	48	48	42	42
R Squared	0.67	0.67	0.67	0.70	0.70
Host Country FE	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes

Notes: The Table shows results for second generation immigrants on the mother's side. *Score Country m* is the average math PISA score of natives (standardized to have mean 0 and standard deviation 1 across all countries participating to the test) in the country of birth of the mother, across all available waves. All specifications control for intercept, students' age (in months), wave fixed effect, a dummy for father's immigrant status and dummies for parents' employment status (full-time employed, part-time employed, not working). *Working* refers to either full-time or part-time employed. *Log GDP Country m*, *Avg Years Edu in m* and *Log Exp per Pupil in m* are respectively the wave-specific contemporaneous log real GDP per capita, the average years of schooling in 2005 of 35- to 45-year-old adults and the log expenditure in 2000 per pupil in secondary schools in the country of birth of the mother. Observations are weighted according to the provided sample weights. Standard errors are clustered by mother's country of origin, and inflated by the estimated measurement error in test scores. * denotes significance at 10%, ** at 5%, *** at 1%.

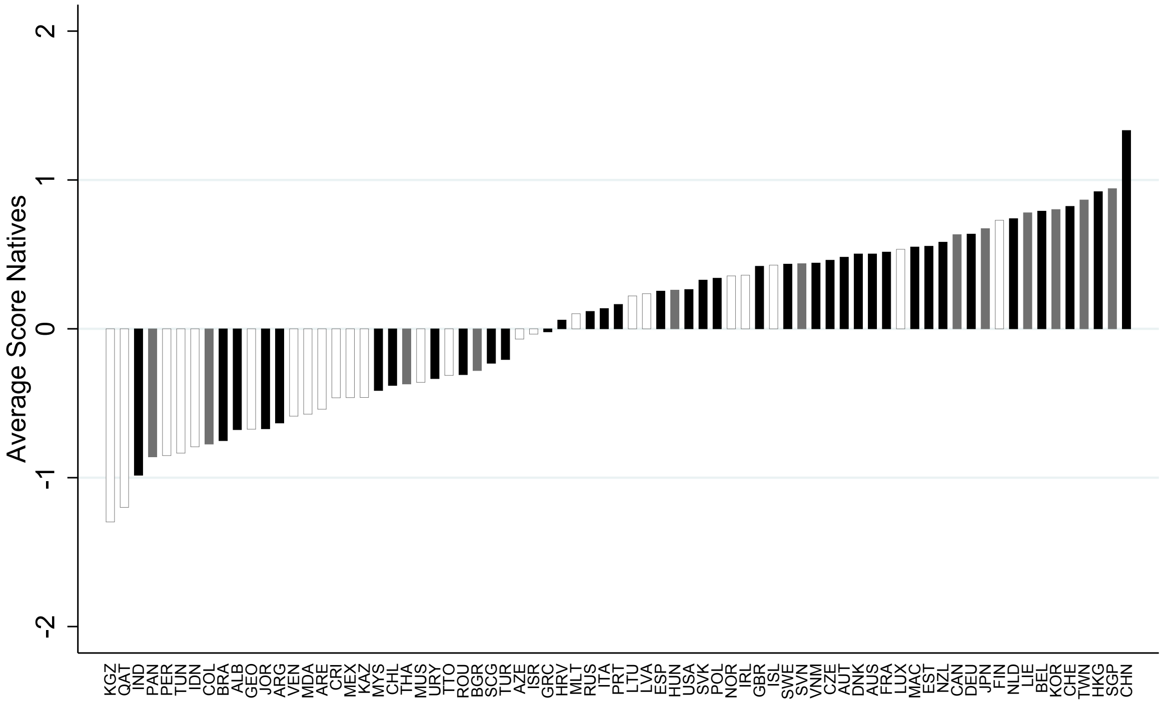
Table 1.10: Country of Origin Characteristics - Cultural Traits

	Dependent variable: Math Test Score					
	[1]	[2]	[3]	[4]	[5]	[6]
Score Country *m*					0.251***	0.158***
					(0.071)	(0.060)
Female	-0.200***	-0.202***	-0.200***	-0.198***	-0.201***	-0.198***
	(0.023)	(0.022)	(0.023)	(0.023)	(0.022)	(0.023)
Father Sec Edu	0.013	0.010	0.012	0.011	0.016	0.014
	(0.023)	(0.023)	(0.023)	(0.023)	(0.022)	(0.022)
Father Ter Edu	0.043	0.044	0.041	0.040	0.047	0.043
	(0.034)	(0.034)	(0.034)	(0.034)	(0.034)	(0.034)
Mother Sec Edu	-0.003	-0.009	-0.007	-0.019	-0.014	-0.020
	(0.041)	(0.040)	(0.040)	(0.035)	(0.038)	(0.034)
Mother Ter Edu	0.002	-0.007	-0.007	-0.021	-0.010	-0.020
	(0.043)	(0.045)	(0.042)	(0.042)	(0.043)	(0.043)
Mother Working × Mother ISEI	0.003***	0.003***	0.003***	0.003***	0.003***	0.003***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Father Working × Father ISEI	0.003***	0.003***	0.003***	0.003***	0.003***	0.003***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Leisure Important in *m*	-0.258***			-0.232***		-0.247***
	(0.088)			(0.057)		(0.070)
Locus of Control in *m*		0.308**		0.440***		0.266**
		(0.140)		(0.087)		(0.112)
Long Term Orientation in *m*			0.445***	0.421***		0.237*
			(0.153)	(0.115)		(0.125)
N	39882	39882	39882	39882	39882	39882
# Country *m*	46	46	46	46	46	46
R Squared	0.67	0.67	0.67	0.68	0.67	0.68
Host Country FE	Yes	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The Table shows results for second generation immigrants on the mother's side. *Score Country m* is the average math PISA score of natives (standardized to have mean 0 and standard deviation 1 across all countries participating to the test) in the country of birth of the mother, across all available waves. All specifications control for intercept, students' age (in months), wave fixed effect, a dummy for father's immigrant status and dummies for parents' employment status (full-time employed, part-time employed, not working). *Working* refers to either full-time or part-time employed. *Leisure Important in m* and *Locus of Control in m* are constructed from answers of natives in the country of birth of the mother to the corresponding questions in the World Value Survey (described in the main text), and are standardized to take mean 0 and standard deviation 1 in the WVS sample. *Long Term Orientation in m* is constructed in Hofstede et al. (2010) and ranges from 0 to 1. Observations are weighted according to the provided sample weights. Standard errors are clustered by mother's country of origin, and inflated by the estimated measurement error in test scores. * denotes significance at 10%, ** at 5%, *** at 1%.

Figures

Figure 1.1: Performance of Native Students across Countries

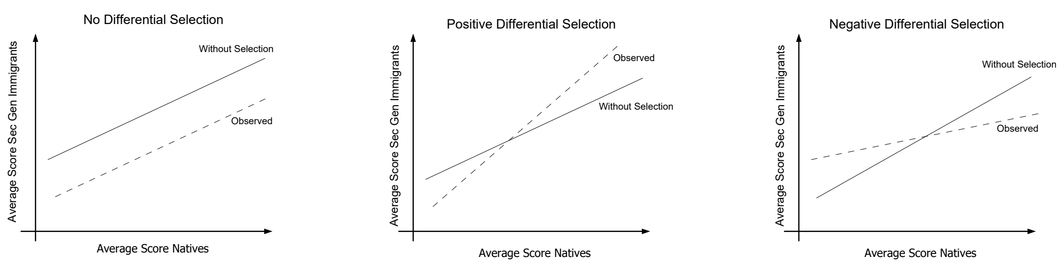


Notes: The height of the bar represents the average PISA score in mathematics for native students. Scores are standardized to have mean 0 and (individual-level) standard deviation 1 across the (pooled, equally weighted) 73 countries participating to at least one wave of the test. Black bars refer to countries in the core sample, grey bars to countries for which we observe at least one second generation immigrant but less than 100 immigrant parents.

Figure 1.2: Performance of Second Generation Immigrants and Natives

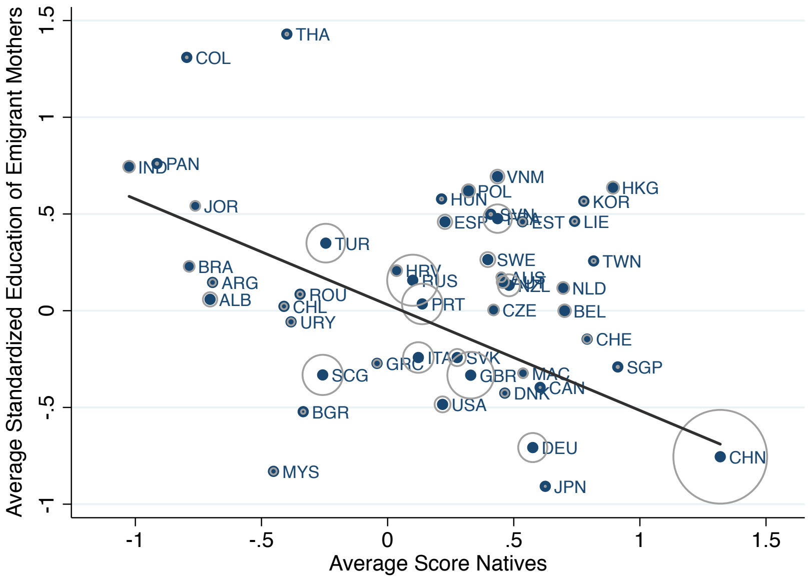
Notes: The left panel plots the average PISA score of second generation immigrants whose mother is from country m against the average math PISA score of natives in country m, for all countries with at least 100 second generation immigrants on the mother's side in the sample. The right panel plots the predicted scores from a regression with individual math scores as dependent variable and fixed effects for mother's country of origin, gender, both parents' education and employment status, father's immigration status and school fixed effects as controls, with all covariates except country of origin fixed effects set at their sample mean and the sample restricted to second generation immigrants on the mother's side. The size of the circles is proportional to the number of second generation immigrants on the mother's side in the sample. The line shows the best (weighted) linear fit.

Figure 1.3: Different Types of Selection



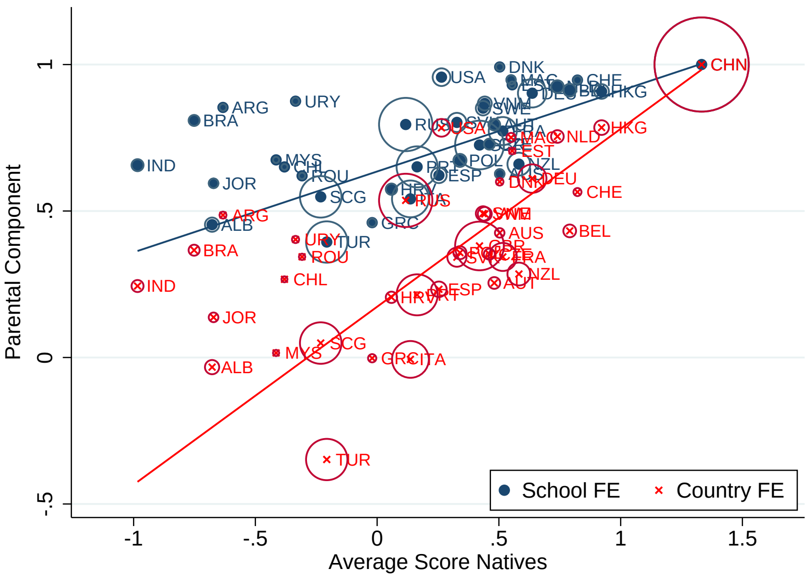
Notes: The Figure represents three possible patterns of emigrant parents' selection on unobservables. The first panel refers to the case where emigrant parents are selected to the same extent across all countries of origin. The second (third) panel refers to the case where emigrant parents from high PISA countries are more positively (negatively) selected.

Figure 1.4: Selection on Parental Education



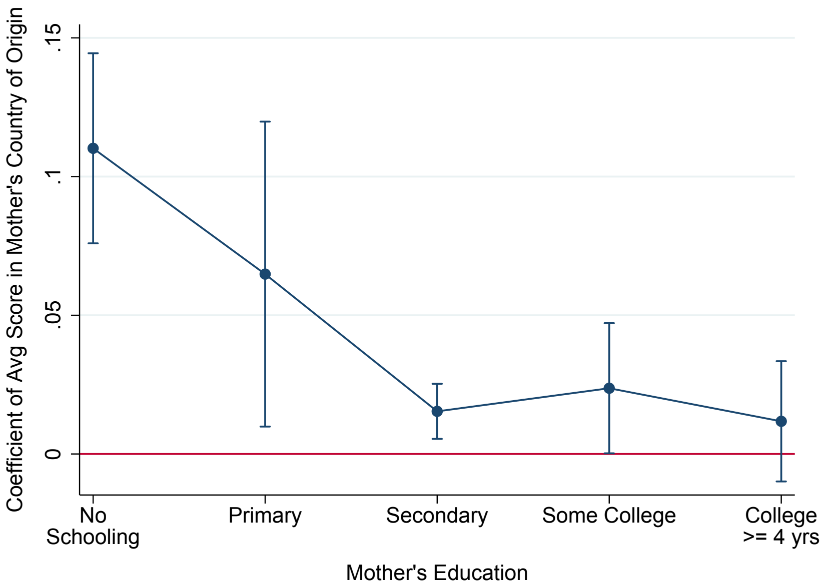
Notes: The Figure plots the average years of schooling of emigrant mothers from country m standardized by the average and the standard deviation of years of schooling of non-emigrant mothers in country m (y-axis) against the average PISA score of native students in country m (x-axis). The sizes of the circles are proportional to the number of emigrant mothers in the sample. The line shows the best (weighted) linear fit.

Figure 1.5: Parental Component



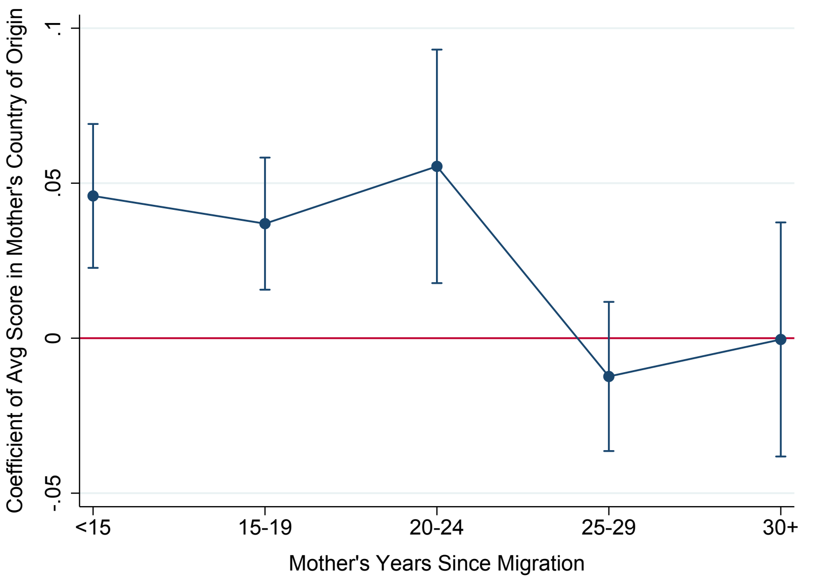
Notes: The Figure plots the estimated parental component ($Parents^c$ as defined in Section 1.6), normalized such that it takes value 1 for China (y-axis) against the average PISA score of natives (x-axis). Only countries with at least 100 emigrant parents in the sample are included. The sizes of the circles are proportional to the number of emigrant parents in the sample. The lines show the best (weighted) linear fits.

Figure 1.6: Heterogeneous Effect with respect to Mother's Education - US Census



Notes: The Figure plots the estimated coefficients and 95% confidence intervals on the interactions between the average PISA score of natives in mother's country of origin and dummies corresponding to mother's educational achievement, with the dependent variable and other controls being the same as in column 5 of Table 1.3. Standard errors are clustered by mother's country of origin.

Figure 1.7: Heterogeneous Effect with respect to Mother's Years Since Migration - US Census



Notes: The Figure plots the estimated coefficients and 95% confidence intervals on the interactions between the average PISA score of natives in mother's country of origin and dummies corresponding to mother's years since migration, with the dependent variable and other controls being the same as in column 5 of Table 1.3. Standard errors are clustered by mother's country of origin.

Chapter 2

Barriers to College Investment and Aggregate Productivity

2.1 Introduction

It is a well known fact that rich families tend to invest more in childrens' human capital compared to poor families. This is true in the US, and in virtually every country where enough data to document these intergenerational patterns are available.[1] This disparity holds for investments at different stages of the life cycle, from early childhood to higher education.

In the US, the relationship between socioeconomic status and college opportunities has become lately object of extensive public debate. Many observers, both academic and non academic, have expressed the concern that the US higher educational system is failing to provide students with a level playing field, where only merit and potential determine the access to better opportunities.[2] While most of the current debate emphasizes equity considerations, these facts might have important consequences also from an efficiency perspective.

In this paper, I investigate the impact of barriers to college investments for low income families on economy wide productivity. If high potential individuals are prevented from accessing to an adequate college education, then they will not be as productive as workers as they could be. The objective of my analysis is to quantify this productivity loss, therefore estimating how much it could be gained from policies that make access to college education more meritocratic.

The fact that individuals from rich families go to college more compared to individuals from poor families is not informative per se of an efficiency loss. Indeed, the burgeoning literature on skills formation emphasizes that the accumulation of human capital is a dynamic process, and that at each stage of the life cycle there are important complementarities between the current stock and the productivity of new investment (Heckman and Cunha, 2007; Cunha et al., 2010). Since children from rich families accumulate more human capital early on, it is natural (and efficient) that *at the college enrollment stage* they will be investing more.[3] In this paper I

[1] See Glewwe and Kremer (2006) for a discussion on developing countries.

[2] See for example Hoxby and Avery (2012) and Paul Tough, "Who Gets to Graduate?", *New York Times Magazine*, 15/05/2014.

[3] The efficiency considerations would be clearly very different if I were to consider the allocation of human

follow the skills formation literature by taking results in test scores administered at the end of high school as measures of the stock of human capital that individuals are endowed with when making decisions on college enrollment. In other words, I ask whether distortions due to family income are important once its effect on early human capital is controlled for.

As I will discuss more in detail below, there are in principle many possible reasons why family income, conditional on ability, correlates with human capital investments on children. A classic explanation is based on credit market imperfections: since future earnings might not be pledged as collateral, individuals from poor families are unable to finance as much education as they would want. Credit constraints however receive mixed empirical support in US data, and several alternative explanations have been proposed.[4] Empirically distinguishing between these alternative frictions is a daunting task, given that many of them are likely to be present at the same time and to interact with each other.

I consider a setting that does not require to take a stand on exactly what is preventing low income families to invest on college as much as rich families do. Instead, I aim to capture the overall effect of this disparity through the reduced form approach introduced by Hsieh and Klenow (2009) in the misallocation literature. In particular, I propose a framework where individuals face different implicit "taxes" when making their college enrollment choices, depending on their family income. These objects should not be literally interpreted as taxes, but as the overall wedges between investment and return to education which might be due to credit constraints, imperfect information or any other friction. I back out these wedges from the structure of the model, and then I implement counterfactual experiments where barriers for low income families are eliminated.

The results of these exercises suggest that the productivity costs stemming from the inequality of college access opportunities might be substantial. Under my baseline parametrization, output and wages would increase by approximately 10% if individuals from low income families were to have the same possibilities of their peers from wealthy families. Most of these gains would come from what I call the "intensive margin" of college investment, which is the amount of human capital accumulated conditional on attending college (as opposed to the "extensive margin", the choice between attending college or not). Therefore, policies aimed to improve efficiency should not aim to achieve large increases of college enrollment rates, but instead to help students from disadvantaged background to attend higher quality schools and make the most of their time there.

This paper speaks to several strands of the literature. First, it is clearly related to the huge literature on the determinants of college enrollment choices, and in particular the disparity in college opportunities between students of different family backgrounds. While this disparity is well documented (Ellwood and Kane, 2000; Hoxby and Avery, 2012) the debate on its determinants is quite open. An explanation explored in the economics literature is that poor families

capital investment at earlier stages. In ongoing work, I am developing a model appropriate for such an analysis. Still, taking a "snapshot" at the college enrollment stage is particularly interesting given the many policies are designed to correct inefficiencies arising at this stage.

[4]This literature is briefly reviewed below.

are subject to borrowing constraints that prevent them to invest in their children's education as much as they would want to (Becker, 1962). The evidence on credit constraints for higher education is rather mixed: Cameron and Heckman (1998), Keane and Wolpin (2001) and Carneiro and Heckman (2002) argue that they are binding for at most a small share of students, while Brown et al. (2012) find that they do play an important role when the different incentives of parent and children are explicitly taken into account. Recent contributions have considered barriers of different nature: Hoxby and Avery (2012) and Carrell and Sacerdote (2013) find that providing information and mentoring are potentially effective ways to induce low income high school students to attend college, while Gorard et al. (2012) emphasize the role of differential attitudes towards college education. Differently from all these papers, my objective here is not to estimate the relative importance of specific barriers to college investment, but instead to evaluate their combined impact on aggregate productivity and wages.

My work is also closely related to a small literature that investigates the macroeconomic costs of human capital misallocation. I draw heavily from the framework proposed by Hsieh et al. (2013), who quantify the contribution of the relaxation of labor market frictions for women and black men to US economic growth in the last few decades. Buera et al. (2011) and Caselli and Gennaioli (2013) study the misallocation of entrepreneurial talent due to credit frictions, while Vollrath (2014) investigates the allocation of human capital across sectors. Differently from these papers, I study the allocation of human capital *investment*, rather than human capital per se.[5] Moreover, I focus on a different source of misallocation, namely the fact that family income shapes the access to education on top of academic ability. The interest in this source of misallocation is shared by Hanushek et al. (2014), who develop a dynamic general equilibrium model to quantify the impact of different policies aimed to relax credit constraints. Differently from their work, I do not restrict my attention to credit constraints, but instead I study a broader range of barriers to college investment.

The paper is structured as follows. In Section 2 I describe the evidence on the disparity of college investment between rich and poor families, both on the extensive and the intensive margin. Section 3 introduces the model, while Section 4 describes the calibration procedure. Section 5 presents the main results, while robustness checks and extensions are left for Section 6. Finally, Section 7 concludes by examining policy implications and avenues for future research.

2.2 Family Income and College Investment

2.2.1 Data

Throughout the paper, I use data from the 1979 wave of the National Longitudinal Survey of the Youth (NLSY79). This dataset provides a nationally representative panel of 12,687 young men and women that were between 14 and 22 years old in 1979. I focus on the main cross-sectional

[5] Hsieh et al. (2013) study the combined effect of frictions relative to educational investment and occupational choice.

sample, and exclude the oversamples of ethnic minorities and disadvantaged individuals. The dataset includes detailed information on education, labor market outcomes and, crucially for my purposes, results of standardized tests designed to measure cognitive and non-cognitive skills that were administered to sample members roughly at the end of high school.

As a measure of the family socioeconomic status, I use total net family income in 1978 and 1979. This should be informative of the resources available to families in the years when college choices are made.[6] I exclude from the sample individuals that do not live either with their parents or at a temporary address (such as a student dorm), since for those family income might not be informative of the actual resources at their disposal when choosing whether to go to college. To soften the concerns about the possible bias arising from short term fluctuations in income, I follow a common practice in the intergenerational mobility literature by taking the simple average of the two years.[7]

Individuals are considered to have attended some college when the highest grade they have completed is 13^{th} or higher.[8] Since the model is not aimed to capture the factors that determine high school completion, I discard all observations relative to high school dropouts (i.e., those individuals whose highest grade completed is 11^{th} or lower); however, results are very similar when these are included.

The main proxy for accumulated human capital that I use in the paper is the result in the Armed Forces Qualifications Test (AFQT). This test is widely used in the labor economics literature as a proxy of cognitive ability, and it is widely recognized to reflect both innate ability and human capital accumulated during childhood (Cascio and Lewis, 2005). I construct the raw AFQT scores by combining the results obtained in different sections of the Armed Services Vocational Aptitude Battery (ASVAB) test, according to the formula documented in NLS (1992). A complication arises from the fact that the test was taken in 1981 by all students in the sample, who at the time were at different grades. In order to clean test scores from the component due to schooling differences in 1981, I adopt the following procedure, which is similar to the one described in Carneiro and Heckman (2002): I divide students in groups according to the highest grade attended in their life, and within each group I take the sum of the constant and residual estimated from a regression of the raw AFQT score on the difference between the grade attended in 1981 and 12.[9,10] The obtained scores are normalized so that they range from 0 to 100.

[6] Strictly speaking, these are the relevant years (mostly) for individuals that are 16 or 17 years old in 1979. I use them for all individual in the samples in order to have a directly comparable measure of family income, which can be used to construct quantiles of interest. Focusing on the younger part of the sample would considerably restrict the number of observations and not alter the major results of the paper (if anything, I find slightly higher counterfactual gains when I limit the sample to individuals that are 16 or 17 years old in 1979; these results are available upon request).

[7] Whenever family income is available for only one of these two years, I include the available measure.

[8] I take the maximum grade completed up to 33 years old, since, as described below, wages are measured at 35.

[9] Dividing by group according to the educational level achieved is necessary since there exist factors (such as family income, as stressed in this paper) that are positively correlated with both schooling at test date (for those with some college education) and ability, so that the effect of schooling on test scores would be overstated in a pooled regression. Instead, when I condition on total education achieved in life, the variation in schooling at test date should depend only on age in 1981.

[10] I also correct for age differences on top of this, even though this adjustment turns out to be mostly inconse-

In the robustness section I also use measures of non-cognitive ability, which the skills formation literature has showed to be important for educational and labor market outcomes. I clean these measures from schooling differences in 1981 using the same procedure outlined above for the AFQT. The exact tests used are described more in detail in the robustness checks section.

The identification strategy adopted in this paper requires a measure of adult labor market income. I construct hourly wages at 35 (or the closest possible alternative) from data on total labor earnings and hours worked available in the NLSY79. Observations for which hours per week are below 10 or above 100 are dropped. Since the measurement of labor earnings is notoriously imprecise at the very bottom and the very top of the distribution, I impute wages corresponding to the 1^{st} and 99^{th} percentiles for individuals below and above these thresholds. In order to net out the effect of characteristics which are not the main focus of the paper, I regress (log) wages on race, gender and age controls, and use the (exponential of the) estimated residuals throughout.

The final sample is composed by 3000 individuals for which I have complete information on family income, education, ability and wages. When I include measures of non-cognitive skills, the sample size drops to 2934. All the summary statistics and regression results reported below make use of the provided sample weights.

2.2.2 The Extensive Margin

In this section I report evidence from the NLSY79 on to the extent to which family income is an important determinant of college enrollment. For this purpose, I split my sample in three groups according to family income terciles. Figure 2.1 shows the share of individuals in each group with at least some college education.[11]

The differences between groups are quite substantial: more than 60% of children from ”High Income” families get some education beyond high school, while the corresponding figure for the ”Low Income” group is just above 40%. As discussed in the introduction, this disparity is not particularly puzzling, given that students coming from rich families are likely to be more prepared for college given that they have attended better schools and in general lived in environments more favourable to human capital accumulation. Indeed, Figure 2.2 documents how these students achieve substantially higher scores in the AFQT test.

In order to understand whether family income represents a barrier for college enrollment on top of its impact on ability, Figure 2.3 breaks down each income group in three subgroups according to test scores terciles.

While there seem to be some differences across income groups, overall the disparity is far from being dramatic. Almost 85% of the individuals in the high income - high AFQT group attend college, while the corresponding figure for the high achievers in the low and middle income group is lower by approximately 10 percentage points. Similar gaps can be noted for

quential.

[11] The college attendance figures reported is this paper are slight higher compared to the ones from other sources (such as, for example, Belley and Lochner (2007)), since here high school dropouts are excluded from the sample. As mentioned above, results do not depend on this sample restriction.

the individuals in the Medium AFQT group, while the differences are even smaller between those that score poorly in the test. On the basis of this and further evidence, Carneiro and Heckman (2002) argue that credit constraints do not seem to play an important role in college enrollment, which is mainly determined by the human capital accumulated during childhood.[12]

2.2.3 The Intensive Margin

On the basis of the evidence presented in the last section, one might be led to conclude that barriers to college investment for low income families are unlikely to be quantitatively important. However, the fact that conditional on ability individuals from different economic backgrounds are almost equally likely to extend their education does not mean that they accumulate the same amount of human capital once they are in college. Indeed, several recent papers have documented that family income is strongly correlated with the quality of college investment, and that this intensive margin is quantitatively important for wages. In this section I briefly review some of these studies, and then I offer some new evidence based on the NLSY79 data.

2.2.3.1 Existing Literature

Even when they attend college, students from low income families appear to pick schools that often are not up to their potential. The issue of "academic undermatch" has been at the center of a small but growing literature in educational economics, which has consistently shown that the problem is pervasive in the US, especially within low income groups and ethnic minorities (see, among the others, Cabrera and La Nasa (2001), Hill and Winston (2010), Pallais and Turner (2006) and Smith et al. (2013)). A particularly enlightening study for my purposes is Smith et al. (2013), which uses nationally representative data to quantify the extent of academic undermatch for students of different socioeconomic groups. According to their definitions, 49.6% of students with a lower socioeconomic status are undermatched, while the corresponding figure for students with a higher socioeconomic status is 34%. The contrast is starker for high achievers: 60% and 50.4% of disadvantaged students who potentially have access to "selective" and "somewhat selective" colleges are undermatched, while the corresponding figures for richer students are 43.3% and 28.7%. A similar message emerges from the work of Hoxby and Avery (2012): the authors document that the majority of low income students who do extremely well in standardized tests do not even apply to selective colleges, and overall follow seemingly inefficient application strategies. In a subsequent paper Hoxby and Turner (2013) argue that a lack of information is at the origin of this puzzling behavior, and that very simple and cost effective policies can lead students to apply to colleges of the appropriate quality (and then succeed in them). Using the data from the 1979 and 1997 waves of the NLSY, Kinsler and Pavan (2011) document that family income strongly affects the quality of the college attended, and that the

[12] Carneiro and Heckman (2002) also show that the relationship between college enrollment and family income is weakened further when factors such as parental education, family structure and place of residence are controlled for.

effect is weaker for the second wave (consistent with the development of more merit based policies over time).

A related finding reported by Hoxby and Avery (2012) and Belley and Lochner (2007) is that students from low income families choose to attend colleges closer to their place of origin. This might reflect a gap in information about better alternatives, as argued by Hoxby and Turner (2013), or more generally the fact that the distance from home embodies a larger cost for low income individuals. Moreover, since they usually come from disadvantaged regions, it is unlikely that they end up in selective institutions.

Substantial disparities in time use during college have also been documented. For example, Keane and Wolpin (2001) and Belley and Lochner (2007) document how poor students are disproportionately more likely to work part time during college, and discuss how this might impact their learning experience.

Are these margins important for productivity? While there is quite convincing evidence on the fact that college quality matters (Black and Smith, 2006; Kinsler and Pavan, 2011), the relevance of many other factors discussed in this section is obviously hard to identify. One advantage of the approach proposed in this paper is that, at the price of some admittedly restrictive assumption, it bypasses such identification problem by relying on the structure of the model to infer the importance of the intensive margin of college investment.

2.2.3.2 New Evidence from the NLSY79

A crucial dimension over which college experiences are highly heterogeneous across US students is given by the type of degree it terminates with. While Figures 2.1 and 2.3 classify as attending college any student who goes beyond the 12th grade, many eventually drop out without obtaining any formal recognition, while others are awarded with bachelors and graduate degrees. Several papers document that the labor market offers a wage premium to individuals with a more advanced degree (Frazis, 1993; Jaeger and Page, 1996; Park, 1999); while it is difficult to disentangle to what extent this is due to a “sheepskin effect” or differential human capital accumulation, it seems uncontroversial that finishing a given degree entails benefits compared to stopping short of it.

In this section I use data from the NLSY79 to document how, conditional on ability, students from low income families that attend some college fare worse in terms of the obtained degree. I do not necessarily wish to claim that the relationship is causal; instead, for the purpose of this paper, it is sufficient to document that there is some aspect associated with family income that correlates with these outcomes even when ability is controlled for.[13]

Table 2.1 shows the estimates from a multinomial logit regression where the dependent variable is the type of degree obtained in college.[14] Only students with some college education are

[13]College quality and resources have been shown to influence whether a given student obtains a degree (Bound and Turner, 2007), so more dropouts for low income students could reflect the fact that they usually attend less selective colleges.

[14]I have experimented with other specifications, such as ordered multinomial logit and probit, obtaining similar results.

included in the sample, and the considered categories are dropout without any degree (omitted), associate degree (including degrees from junior colleges), bachelor, graduate degree (including Masters, PhDs and professional degrees) and other degree. The regressors include (log) family income, ability (as measured by the AFQT score) and various demographic controls.[15] A positive coefficient on (log) family income implies that, conditional on ability, socioeconomic background is positively associated to the probability of obtaining a given degree. This is true for both bachelor and graduate degrees, while not significantly so for associate and other degrees. Not surprisingly, ability is positively related to the probability of obtaining all types of degrees (relative to not getting any).

In order to interpret the magnitude of the results, Table 2.8 displays the predicted probabilities of obtaining each type of degree (conditional on attending college) for individuals belonging to different terciles of the family income distribution, with the other controls evaluated at their sample average. It emerges that a student from the lowest family income tercile is approximately 10 percentage points more likely to dropout compared to one from the highest income tercile with the same (average) level of ability; by contrast, the latter is 11 and 4 percentage points more likely to obtain a bachelor and graduate degree compared to the former.

2.3 Model

The economy is populated by a unitary mass of agents who have just graduated from high school. They are heterogeneous along two dimensions: ability z and family income y. Ability here refers to the stock of human capital they have accumulated so far, which might be a composite of innate skills and previous investments. More specifically, agents belong to 1 of 3 ability groups and to 1 of 3 income groups, which map in the test scores and family income terciles considered in section 2; there are therefore 9 income - ability groups in the economy.[16]

The model is static. Agents choose whether they want to work as low skilled workers or go to college and work as high skilled. Moreover, if they go to college they have a choice on how much to invest in their university education; in particular, they acquire a certain number of "educational units" e. This is meant to capture the intensive margin of college investment discussed in the previous section: a higher e corresponds to attending a more selective school, putting more effort in the learning experience, participating to extra-curricular activities and in general taking advantage of every factor that affects the amount of human capital that is accumulated during college. I bundle all these aspects together because, as discussed above, it is very difficult to empirically separate between them, and, moreover, there are possibly many other important factors which are completely unobservable to the econometrician. As will become clear later, my strategy here is to use the structure of the model to back out the overall importance of these margins.

Going to college involves giving up a fraction s of the wage. This might be interpreted

[15]Results are very similar when measures of non cognitive skills are included as well.

[16]This classification in 9 groups is convenient but clearly arbitrary. In Section 2.6.3 I consider deviations from this modeling choice.

as a time cost: if an agent chooses to go to college, he will work only a fraction $1 - s$ of his lifetime, while by opting for the low skill sector he can start working immediately.[17] The cost s is increasing in the amount of efficiency units that an individual wants to acquire. The rate at which individuals can convert s into e depends on ability and family income through an exogenous wedge $\tau(y, z)$; in particular

$$s = e(1 + \tau(y, z)) \tag{2.1}$$

The fact that the wedge $\tau(y, z)$ is a function of ability reflects the finding in the skill formation technology literature that the existing stock of human capital directly affects the productivity of new investments (Heckman and Cunha, 2007; Cunha et al., 2010). This might be because higher cognitive skills facilitate learning, or simply because new knowledge is more productive when built on a stronger basis. Therefore, I expect $\tau(y, z)$ to be decreasing in z.[18]

Barriers to college investment related to family income are captured by the fact that $\tau(y, z)$ is potentially a function of y. If $\tau(y, z)$ is decreasing in y, a student with high family income will be more efficient in accumulating educational units compared to one with low family income. To what extent this is the case is the main object of interest of the paper.

Agents make their educational choice to maximize consumption, which is simply equal to the wage (net of college cost if they choose to attend it). If agent i chooses to work in the low skill sector, he obtains a wage equal to

$$w_L(i) = r_L z(i)^\alpha \varepsilon_L(i) \tag{2.2}$$

where r_L is the price of an efficiency unit in the low skill sector and $\varepsilon_L(i)$ is an idiosyncratic shock. In the spirit of a Roy (1951) model, $\varepsilon_L(i)$ embodies all unobservable factors that make an individual more or less productive in a certain sector; this shock represents the only source of heterogeneity between agents in the same income - ability group.

If agent i chooses instead to work in the high skill sector, he obtains a wage equal to

$$w_H(i) = r_H e(i)^\eta \varepsilon_H(i) \tag{2.3}$$

where r_H is the price of an efficiency unit in the high skill sector, $\varepsilon_H(i)$ is an idiosyncratic shock and $e(i)$ is the amount of education units agent i acquires in college. This wage depends on $z(i)$ indirectly through the impact that the latter has on the choice of $e(i)$.[19]

The optimal amount of educational units acquired by an individual going to college is given

[17] However, s does not have to be necessarily interpreted as a time cost. Any cost that enters proportionally to the wage would fit in this setting. Having only proportional costs greatly simplifies the analysis.

[18] Under the time cost interpretation of s, this means that a student with low ability that wants to achieve the same number of educational units of a student with high ability will have to invest more time in college.

[19] Any direct impact of $z(i)$ on the number of efficiency units supplied to the high skill sector is also captured by $\tau(y, z)$. Such an impact is difficult to (separately) identify from the data.

by the solution of this simple problem

$$\max_{e(i)} \left[1 - e\left(1 + \tau\left(y(i), z(i)\right)\right)\right] r_H e(i)^\eta \varepsilon_H(i) \tag{2.4}$$

and is given by

$$e^*(i) = \frac{\eta}{(1+\eta)(1+\tau(y(i), z(i)))} \tag{2.5}$$

From (2.5) it emerges that the optimal amount of educational units does not depend on the realization of the idiosyncratic shock, and is therefore the same for every individual belonging to the same income - ability group. This result greatly simplifies the inference problem, given that it requires me to back out only one object for each income - ability group.

Agent i anticipates how much he will be able to invest in college before deciding whether to enroll or not. Let $S(i)$ be a dummy variable equal to 1 if i goes to college and to 0 if he does not. Plugging $e^*(i)$ from (2.5) in the objective function given in (2.4), I obtain that consumption as a function of the educational choice is

$$c(i) = \begin{cases} r_L z(i)^\alpha \varepsilon_L(i) \text{ if } S(i) = 0 \\ \frac{\bar{\eta} r_H}{(1+\tau(y(i), z(i)))^\eta} \varepsilon_H(i) \text{ if } S(i) = 1 \end{cases}$$

where $\bar{\eta} = \frac{\eta^\eta}{(1+\eta)^{1+\eta}}$. The choice between going and not going to college takes the form of a standard discrete choice problem, where the value of the two alternatives is proportional to two unobservable shocks. I follow a common practice in discrete choice econometrics by assuming that these shocks are extracted from two independent Frechet distributions, with cumulative density functions given by

$$F(\epsilon_L) = e^{-\epsilon_L^{-\theta}}$$

$$F(\epsilon_H) = e^{-\epsilon_H^{-\theta}}$$

where θ is a parameter inversely related to the variance of the shock.[20] Under this distributional assumption, it is straightforward to show that the probability that agent i with family income y and ability z goes to college is[21]

$$P[S(y,z) = 1] = \frac{(\bar{\eta} r_H)^\theta}{(\bar{\eta} r_H)^\theta + (r_L z^\alpha (1 + \tau(y,z))^\eta)^\theta} \tag{2.6}$$

By the law of large numbers, this also represents the share of individuals in the (y, z) group enrolling in college. From (2.6) it is immediate to see that this share is decreasing in $\tau(y, z)$: the more inefficient a group is in accumulating educational units, the lower the share of individuals in that group that choose to attend college.

Applying again the law of large numbers, the average wage for individuals in the (y, z) group employed in the low skilled sector is given by

[20] The independence assumption can be relaxed without particular complications.

[21] See the Appendix for a complete derivation.

$$\begin{aligned} w_L(y,z) &= r_L z^{\alpha} E[\epsilon_L(i)|S(y,z)=0] \\ &= \left[\left(\frac{\bar{\eta} r_H}{(1+\tau(y,z))^{\eta}}\right)^{\theta} + (r_L z^{\alpha})^{\theta}\right]^{\frac{1}{\theta}} \Gamma\left(1-\frac{1}{\theta}\right) \end{aligned} \tag{2.7}$$

while the average wages of those employed in the high skill sector is

$$\begin{aligned} w_H(y,z) &= \frac{\bar{\eta}(1+\eta) r_H}{(1+\tau(y,z))^{\eta}} E[\epsilon_H(i)|S(y,z)=1] \\ &= (1+\eta)\left[\left(\frac{\bar{\eta} r_H}{(1+\tau(y,z))^{\eta}}\right)^{\theta} + (r_L z^{\alpha})^{\theta}\right]^{\frac{1}{\theta}} \Gamma\left(1-\frac{1}{\theta}\right) \end{aligned} \tag{2.8}$$

where $\Gamma(.)$ is the gamma function. These results follow from the standard extreme value property of the Frechet distributions; a complete derivation is relegated to the Appendix. Plugging (2.8) in (2.6), the share of individuals in the (y,z) group attending college can be written as

$$P[S(y,z)=1] = \left[\frac{(1+\eta)(\bar{\eta} r_H)}{w_H(y,z)(1+\tau(y,z))^{\eta}}\right]^{\theta}$$

Therefore, for any pair of income - ability groups (y,z) and $(\hat{y},\hat{z})$ we have that

$$\frac{1+\tau(y,z)}{1+\tau(\hat{y},\hat{z})} = \left[\frac{w_H(\hat{y},\hat{z})}{w_H(y,z)}\left(\frac{P[S(\hat{y},\hat{z})=1]}{P[S(y,z)=1]}\right)^{\frac{1}{\theta}}\right]^{\frac{1}{\eta}} \tag{2.9}$$

Equation (2.9) relates the relative friction faced by two income - ability groups to the relative average wage and the relative share of those attending college. Since both wages and college enrollment decisions are observable in the data, I can use (2.9) to back out the relative friction faced by each group (conditional on setting a value for the parameters η and θ; more on this below). Through the lens of the model, a group is inferred to face a large barrier to college investment whenever a few members of that group go to college (low $P[S(y,z)=1]$), and the ones who do earn a low wage afterwards (low $w_H(y,z)$). The overall importance of the (partially) unobservable intensive margin of college investment can therefore be inferred from data on wages: a low investment implies that few educational units were accumulated in college, and this is reflected in a low productivity in the labor market.

The model is closed by the postulation of an aggregate production function that combines the efficiency units supplied in the low and high skill sector to produce an homogeneous good. I assume that the production function takes the standard CES form,

$$Y = A\left[L^{\rho} + BH^{\rho}\right]^{\frac{1}{\rho}} \tag{2.10}$$

where L and H are the total efficiency units supplied in the two sectors and $\frac{1}{1-\rho}$ is the elasticity of substitution. The equilibrium definition is standard; see the Appendix for a formal statement.

2.4 Calibration

In order to perform the counterfactual analysis, I need to set a value for the following parameters: α, η, θ, ρ and B.[22] Moreover, equation (2.9) only provides me with the relative $\tau(y, z)$'s across groups: in order to back out the absolute value of these frictions, I need to impose a normalization on one of them. In this section I discuss the calibration procedure and evaluate its success in matching quantities which are not directly targeted.

First of all, I normalize $\tau(y, z) = 0$ for individuals in the top tercile for both family income and ability. This amounts to saying that these individuals do not face frictions in the acquisition of educational units; since the counterfactual analysis will consist in removing differences in frictions between groups, this is without loss of generality.

I follow Hsieh et al. (2013) in mapping θ to the variance of the residual wages (after the contribution of ability and schooling has been washed out). In particular, it can be verified that within each income - ability group wages follow a Frechet distribution with shape parameter θ. As a consequence of this, the (squared) coefficient of variation of residual wages is equal to

$$\frac{Var\left[w(y,z)|y,z\right]}{\left(\mathbb{E}\left[w(y,z)|y,z\right]\right)^2} = \frac{\Gamma(1-\frac{2}{\theta})}{\left(\Gamma(1-\frac{1}{\theta})\right)^2} - 1 \tag{2.11}$$

In order to construct a measure of residual wages, I take the exponential of the residuals from a regression of log wages on income - ability group dummies, schooling attainment and experience. I then compute the mean and the variance of the exponential of such residuals, and I solve equation (2.11) numerically. The resulting value for θ is 3.27, which is close to the one used by Hsieh et al. (2013).

I estimate α and η using the structure that the model imposes on wages. In particular, the average wage conditional on family income, ability and college enrollment choice is given by

$$\mathbb{E}\left[w(i)|y(i),z(i),S(i)\right] = (1+\eta S(i))\left[\left(\frac{\bar{\eta} r_H}{(1+\tau(y(i),z(i)))^{\eta}}\right)^{\theta} + \left(r_L z(i)^{\alpha}\right)^{\theta}\right]^{\frac{1}{\theta}} \Gamma\left(1-\frac{1}{\theta}\right) \tag{2.12}$$

I plug in (2.12) the values of θ and the $\tau(y, z)$'s obtained from (2.11) and (2.9), and I estimate α and η by non-linear least squares. In order to be consistent with the assumptions of the model, I use only the between groups variation in z when estimating (2.12): in other words, for each individual I set z equal to the mean of the ability tercile to which he belongs.[23] The resulting value for η is 0.32, while the estimate of α is small and not significantly different from zero; I therefore set it equal to zero and examine the effect of different values in the robustness checks.[24]

[22] The value of A is not needed to compute the counterfactual percental change in output and wages.

[23] Section 2.6.3 discusses the importance of this feature of the model for the counterfactual results.

[24] According to this result, the widely documented positive correlation between wages and cognitive test scores for high school educated workers would be entirely due to a selection effect: high cognitive ability makes getting a college education easy, therefore a worker with such an attribute that chooses to work in the low skill sector must have some unobserved comparative advantage in that sector, which is responsible for his high wage. In the robustness checks I document that the results are essentially unchanged when I instead attribute all the positive

I set ρ so that the elasticity of substitution between low and high skilled workers is 1.4, as estimated by Ciccone and Peri (2006). Finally, B is set to match the overall share of individuals with some education beyond high school, which is equal to 56%.

2.4.1 Model Fit

Before moving to the results, it is useful to evaluate how closely the model matches quantities which are not directly targeted by the calibration procedure.

Table 2.8 shows data on college attendance probabilities by income - ability groups, and the corresponding figures implied by the model. In my calibration, $\tau(y, z)$'s are picked so that the relative college shares are consistent with the data, according to equation (2.9). However, nothing ensures that the absolute figures are matched as well.

The model does a reasonable job in capturing the patterns of college enrollment across groups. While the fit is very good for the low income groups, the model somewhat overstates the differences between middle and high income groups.

Table 2.8 reports wage ratios relative to a few groups of interest. The main calibrated parameter that determines relative wages is η, which in the model corresponds to the (constant across groups) college premium. Not too surprisingly then, through the choice of this parameter it is possible to generate an average college premium quite close to the one observed in the data, as shown in the first row of Table 2.8. The model is also quite successful in replicating the wage gaps between the first and third terciles of family income, while it overstates the ability premium. The second and third panel of Table 2.8 display ability and family income gaps within educational groups. While the fit is reasonably good for the income gaps, the model tends once again to exaggerate the importance of ability for wages, especially for high school educated workers. Recall that α, the parameter that controls the "productive" role of ability in the low skill sector, is estimated to be 0, and therefore the entire gap is explained by a selection effect (see footnote 24). The fact that this selection effect is so strong might suggest that the model is slightly overstating the extent of comparative advantage dispersion in the low skill sector (and, to a lower extent, in the high skill sector as well).

With the caveats described above, the model overall does a reasonable job in matching key facts from the data, and therefore can be used informatively for counterfactual analyses.

2.5 Results

Figure 2.4 shows the $\tau(y, z)$'s backed out from (2.9). By construction, members of the high income high ability group have $\tau(y, z) = 0$, so that the other $\tau(y, z)$'s should be interpreted as differences with this benchmark group of privileged individuals.

First of all, within each income group frictions are decreasing in ability at the end of high school. This was to be expected, given that ability is known to be an important input of the

correlation to a productivity enhancing role for cognitive ability (as estimated by OLS).

human capital production function during college: more able individuals are just more efficient at accumulating further knowledge. Moreover, for any ability level, family income is negatively associated with the magnitude of the estimated barriers: through the lens of the model, this reflects the inequality of opportunity of access to college resources. It is interesting to note that individuals with high ability belonging to the intermediate income group face slightly bigger frictions compared to individuals with the same ability level belonging to the bottom income group. This could be due to the fact that governmental support (through reduced tuition, student loans..) is mainly targeted to well performing students at the bottom of the income distribution (Hoxby and Avery, 2012), who therefore face a lower effective price of schooling compared to their peers from the middle class.

Given that these $\tau(y, z)$ are reduced form and fundamentally a-theoretical objects, their overall importance is difficult to evaluate just by staring at Figure 2.4. A more fruitful exercise is instead to use the model to have a sense of the possible economic gains that could be achieved through their elimination: this will provide us with a readily interpretable measure of the efficiency losses stemming from the inequality of educational opportunities.

What is the relevant counterfactual? As discussed above, the frictions displayed in Figure 2.4 reflect in part "technological" barriers due to differences in ability, and therefore a complete elimination of those is not something that policy would achieve easily, at least in the short run.[25] Instead, the counterfactual of interest is getting rid of the variation in the barriers which is due to family income, as shown in Figure 2.5. This corresponds to a world where all individuals belonging to the same ability group have ex ante the same opportunites to attend college, while family background is not a relevant factor for educational choices (after having accounted for its correlation with academic ability).

This counterfactual exercise measures the overall gains stemming from a more meritocratic allocation of college investment, but obviously it does not say much on which (if any) policy should be implemented in order to reap these benefits. The main objective of this analysis is to understand whether potential gains are large, and therefore to what extent it is worthwhile to explore in detail the effectiveness of different policy tools that might achieve some of them.

Table 2.8 displays the counterfactual responses of output, wages and college enrollment share should such a reduction in barriers to college investment be implemented. Gains are substantial: output increases by approximately 11%, while wages in the low and high skill sector increase by 12% and 9% respectively. These gains are due to both an overall increase in investment in educational units and to a better allocation of the existing investment, in the sense that more individuals with high ability at the end of high school are able to go to college (section 2.6.4 discusses the relative importance of these two forces). It is interesting to note how these large increases in productive efficiency would not require a large expansion of the pool of the college enrollment rate, which goes up only by 2 percentage points (from a baseline of 56%). Intuitively, decreasing returns for high skilled workers in the production function puts a limit on how much college enrollment can increase, and as students from disadvantaged backgrounds

[25] Of course there might be sensible policies aimed to reduce the disparity in ability at the end of high school, but this framework is not well suited to analyze them.

manage to enroll, others that are currently attending would find it convenient to do with the high school diploma in the counterfactual.[26]

An additional result that emerges from Table 2.8 is that the increase in average wages is highest for the low skill sector, so that the elimination of barriers would reduce the skill premium by approximately 3 percentage points.[27] In order to better understand this finding, Table 2.8 decomposes the variation in average wages between changes in the price of an efficiency unit (r_L and r_H in the model) and the average number of efficiency units provided to each sector. The elimination of barriers to college investment causes a large increase in the productivity of high skilled workers: more students of high ability can now access higher education and accumulate a large number of educational units. This implies, through a standard general equilibrium effect, a decrease of 5% in the price of one educational unit supplied to the high skill sector; decreasing returns in the production function therefore mitigate the increase in wages for college educated workers. Conversely, in the low skill sector there is a modest increase in the number of efficiency units per worker, accompanied by a substantial increase in the price of one efficiency units, due to the complementarity between low and high skill labor.

Table 2.8 documents how the average wages for different income - ability groups are affected by the counterfactual experiment. Intuitively, individuals from low and middle income families reap most of the benefits stemming from this experiment, with increases in wages of 21% and 19% respectively. Within these groups, students of all ability levels see their wages increase, and particularly so the ones at the middle and top of the spectrum. The only "losers" from the abolition of barriers to college investments are those coming from a high income family, who see their previous advantage in college access vanish. Perhaps counter intuitively, the low ability members of this group are relatively better off: this is due to the fact that many of them were opting for the low skill sector anyway, and now they enjoy a higher price per efficiency unit because of the general equilibrium effect.

What is the relative importance of the extensive and intensive margin of college investment for these results? The second column of Table 2.8 provides an answer to this question. Here I display the counterfactual changes in average wages that take place when only the college attendance rate (extensive margin) is allowed to respond to the reduction in barriers as in the counterfactual, while the number of efficiency units per worker and their prices are kept constant.[28]

[26]The fact that individuals from low income families that start to attend college in the counterfactual have high ability on average amplifies the extent to which this general equilibrium effect kicks in, since they provide a large number of efficiency units to the high skill sector.

[27]This does not strictly correspond to the college premium usually estimated in the literature, since here I label as skilled all individuals with some education beyond high school (and not only college graduates).

[28]The percentage change in average wages for the (y, z) group can be written as

$$\frac{\Delta w(y,z)}{w(y,z)} = \frac{\Delta w_L(y,z) + \Delta P[S(y,z)=1](w_H(y,z) - w_L(y,z)) + P_c[S(y,z)=1](\Delta w_H(y,z) - \Delta w_L(y,z))}{w(y,z)}$$

where the lower script c denotes counterfactual quantities. The contribution of the extensive margin is defined as

$$\frac{\Delta P[S(y,z)=1](w_H(y,z) - w_L(y,z))}{w(y,z)}$$

It emerges that the extensive margin plays only a minor role in the counterfactual increase of wages for individuals in the low and middle income groups, and approximately 85% of the wage gains are accounted for by the intensive margin.[29] This result confirms the importance of not ignoring the intensive margin when discussing the disparity in educational opportunities between students of different backgrounds.

2.6 Robustness and Extensions

2.6.1 Non Cognitive Skills

The results exposed in the previous section might be misleading if scores in the AFQT test did not reflect properly the actual differences in the productivity of college investment between students coming from high and low income families. In particular, the gains from the counterfactual experiment would be overestimated if there existed some component of academic ability which (i) is not fully captured by the AFQT test, (ii) is nevertheless important for productivity in college and (iii) is more abundant in children coming from rich families.

A natural candidate is given by a set of personal traits that is commonly summarized as *non-cognitive skills*, and includes motivation, persistence, self-esteem and self-control. There is growing evidence that non-cognitive skills matter at least as much as cognitive ones for schooling performance and labor market outcomes, both in the economics (Rubinstein and Heckman, 2001; Heckman et al., 2006; Cunha et al., 2010) and in the psychology literatures (Wolfe and Johnson, 1995; Duckworth and Seligman, 2005). If the $\tau(y, z)$'s backed out from the data capture in part differences in non-cognitive skills, then the counterfactual experiment considered in the previous section is not appropriate, given that differences in skill endowments is not something that policy can easily address at the college enrollment stage.

In this section I construct an alternative measure of ability that takes into account non-cognitive skills, and I describe the results of the corresponding counterfactual experiment. I use two proxies of non-cognitive skills which are commonly employed in the literature and both available in the NLSY79: the Rosenberg Self-Esteem Scale and the Rotter Locus of Control Scale. The Rosenberg Self-Esteem Scale measures an individual's degree of approval or disapproval toward himself; it is composed of 10 statements (such as "*I feel that I have a number of good qualities*", or "*I take a positive attitude toward myself*") to which respondents are asked to agree or disagree. The Rotter Locus of Control Scale measures the extent to which individuals believe to have control over their lives, as opposed to the extent to which external factors (such as luck) determine their personal outcomes; it is composed by four pairs of statements (such as "*What happens to me is my own doing*" versus "*Sometimes I feel that I don't have enough control over the direction my life is taking*"), between which respondents choose the one closer to their opinion. Both measures are converted to the same scale of the AFQT test (from 0 to

[29] Not surprisingly, the extensive margin is important for students from high income families, for whom the counterfactual experiment does not imply any change in barriers to investment.

100).

Before setting up the alternative counterfactual experiment, it is worthwhile to check whether it is the case that individuals with a similar AFQT score and different family backgrounds have very different non-cognitive skills. Table 2.8 displays the average scores in the Rosenberg and Rotter Scales for each family income - ability group, where ability is measured by the AFQT as in the previous sections.

There do not seem to be large differences between income groups for people with a similar AFQT score. While high income groups do obtain slightly higher scores, the gaps are very small, suggesting that the cognitive test does not systematically underestimates differences in academic ability across different economic backgrounds.

In order to investigate the importance of non-cognitive skills for the counterfactual results, I construct a new measure of ability that combines the AFQT, Rosenberg and Rotter scores, and perform the whole analysis using terciles of this new measure (instead of the AFQT alone). How to evaluate the relative importance of the 3 tests? I adopt the following approach: I estimate the elasticities of wages with respect to the 3 test scores, α_{AFQT}, $\alpha_{Rosenberg}$ and α_{Rotter}, from a log-wage regression with schooling and experience as additional controls. Then I combine the 3 measures with a simple Cobb-Douglas aggregator,

$$\tilde{z} = (z_{AFQT})^{\alpha_{AFQT}} (z_{Rosenberg})^{\alpha_{Rosenberg}} (z_{Rotter})^{\alpha_{Rotter}}$$

and I use the terciles of $\tilde{z}$ to construct the new income-ability groups. The estimated elasticities are $\hat{\alpha}_{AFQT} = 0.25$, $\hat{\alpha}_{Rosenberg} = 0.17$ and $\hat{\alpha}_{Rotter} = 0.02$.[30] The model is then re-calibrated using the same procedure described above.

Table 2.8 shows the counterfactual results when using this more comprehensive measure of ability. The order of magnitude of the impact on output, wages and college enrollment rate is very similar to the one of Table 2.8. If anything, the gains are slightly higher when non-cognitive ability is taken into account: this reflects the fact that non-cognitive skills are more equally distributed across family income groups, resulting in higher estimated barriers. Overall, the inclusion of non-cognitive skills seems unlikely to affect the main conclusion from the counterfactual exercise.

2.6.2 The Productive Role of Ability

According to the calibration procedure adopted in this paper, ability does not seem to play a direct role in affecting the efficiency units supplied to the low skill sector, and the positive correlation with wages observed in the data is entirely due to a selection effect (see footnote 24). Since this selection effect is, to my knowledge, unexplored in the literature, one might wonder how much the large counterfactual increases in output and wage hinge on this feature of the model.

To address this issue, in this section I examine the sensitivity of the results to the value of α.

[30] The standardized coefficients are $\hat{\alpha}^S_{AFQT} = 0.12$, $\hat{\alpha}^S_{Rosenberg} = 0.07$ and $\hat{\alpha}^S_{Rotter} = 0.01$.

In particular, I consider the opposite extreme to what it emerges from the baseline calibration, by attributing all the observed positive correlation to a direct productive role of ability. I estimate a log wage regression on ability and experience controls (restricting the sample to high school graduates), and I use the estimated elasticity with respect to ability to calibrate α. I do so for both the baseline measure and the one that includes non-cognitive skills: the resulting coefficients are 0.3 and 0.94.[31]

Table 2.8 shows the counterfactual results when these values for α are used in the calibration, for both measures of ability. Output, wages and college attendance rate increase by a very similar amount to the one displayed in Tables 2.8 and 2.8. Therefore, results do not seem very sensitive to the value of α.

2.6.3 Functional Form of $\tau(y, z)$

The model postulates that $\tau(y, z)$ varies only across family income and ability terciles, and not within them. This assumption has the merit of allowing a transparent exposition of the underlying identification strategy, but it is clearly ad hoc and potentially restrictive. In particular, an obvious concern derives from the fact that the within group variation in family income and ability is ignored: if, within each group, students from rich families are also the ones with higher ability, the counterfactual exercise might be overstating the extent to which barriers could possibly be removed. In order to have a first check on whether this is likely to be a substantial problem, Figure 2.6 displays for each group the scatter plot of (log) ability and (log) family income, and the linear best fit. While some of the relationships appear to be upward sloping (particularly so for the low income - low ability group), most of them look pretty flat, suggesting that the 9 groups capture most of the relevant variation.

Nevertheless, it is important to investigate to what extent the counterfactual results depend on this restrictive specification. A natural way to do so within the current framework is to consider a finer classification of ability: if the within group correlation of family income and ability leads to overstate the counterfactual gains, reducing the extent to which ability varies within groups should alleviate this bias and give more reliable results. Of course there is a trade-off between the number of ability quantiles considered and the sample size within each group: in particular, smaller groups imply noisier estimates of the average wages for low and high skilled workers.

Table 2.8 shows how the counterfactual results change when the number of ability quantiles used varies.[32] The magnitude of the results is essentially constant across specifications, with output gains ranging from 9.7% to 10.8%. Allowing more variation in ability within family income groups does not seem to have appreciable impact on the implied counterfactual gains.

As an additional test, I consider a version of the model in which $\tau(y, z)$ is assumed to depend

[31]The standardized coefficients are 0.12 and 0.14.

[32]When 7 ability quantiles are used, average wages for skilled workers are estimated from as little as 11 observations in one income-ability group.

linearly on ability. In particular, I assume that

$$\tau(y, z) = \alpha_y + \beta_y z \tag{2.13}$$

where both the intercept α_y and the slope $\beta_y z$ parameters are specific to each family income group. While obviously imposing a specific functional form, this specification has the the merit of exploiting the whole variation in z observed in the data.[33]

I estimate the parameters of (2.13) directly from the wage equation (2.12) by NLS. Figure 2.7 displays the fitted $\tau(y, z)$'s for all the individuals in the sample. Similarly to what emerges from the baseline version of the model, there seem to be significant gaps in educational opportunities between individuals belonging to different family income groups and with a given level of ability. These gaps decline in magnitude throughout the ability distribution, reflecting the fact that an increase in ability brings a relatively large benefit in terms of college opportunities to students from disadvantaged backgrounds. The relevant counterfactual experiment involves once again eliminating all the variation in $\tau(y, z)$ brought about by y, while keeping the part stemming from z, as shown in Figure 2.8.

Setting up the counterfactual in this version of the model involves some additional difficulties. To see why, recall that in the baseline model the probability of attending college (2.6) and average wages (2.7) and (2.8) conditional on y and z are directly mapped to the data through a "law of large numbers" type of argument; such a logic does not apply here, given that I observe at most one individual for any relevant pair (y, z). While a full description of the procedure I use is relegated to the Appendix, I provide here a short outline of the key steps.

First, I estimate r_L and r_H, along with the parameters of (2.13), directly from (2.12) by NLS. From the coefficients' estimates and data on wages, I back out the implied $\varepsilon_L(i)$ for all individuals not attending college and $\varepsilon_H(i)$ for those attending college, and compute L and H by summing up all the efficiency units supplied to the two sectors. To perform a counterfactual analysis, I need however to impute a value also for the shock relative to the occupation that is not chosen by the individuals in the sample, in order to be able to evaluate whether they would be attending college under the new set of counterfactual parameters. Given that obviously I do not observe any wage for the unchosen occupation, all that I know is that these shocks are distributed according to a truncated Frechet distribution, where the truncation point is individual specific and represents the highest possible realization consistent with the observed educational choice.

To make progress, I simulate the unobservable shocks from these individual specific distributions. In particular, I draw 1000 sets of shocks, and compute the counterfactual outcomes for each realization. Table 2.8 reports the 5^{th} and 95^{th} of the simulated counterfactual changes in output, wages and college attendance share. The range of results is rather narrow, and the gains, while slightly smaller, are still of comparable magnitude to the baseline counterfactual.

[33] In principle one could go beyond this and allow $\tau(y, z)$ to be a flexible polynomial in z and y. In practice, such a specification is difficult to implement because of multicollinearity problems in the estimation of the polynomial's parameters.

Overall, the results in this section suggest that the within group correlation between family income and ability is unlikely to be quantitatively important for the conclusions of the paper.

2.6.4 Misallocation of College Investment

In the previous sections I have documented how large gains in output and wages can be achieved by expanding college opportunities for students from low and middle income families. These gains come from two different sources: on one hand the overall investment in college education increases, on the other hand existing resources are allocated more efficiently, i.e. to students that are better equipped to make the most out of them. It is natural to ask what is the relative importance of these two forces, since they have different implications on the feasibility of policies aimed to achieve these gains: while improving the allocation of existing resources might be relatively cost-effective, there might be economical and physical limitations to the extent to which these resources can be expanded.

It is unfortunately difficult to pin down the exact answer to this question within the current framework. The reduced form modeling strategy that I adopt in this paper has the merit to allow a quantification of a very broad notion of college investment, but this comes at the cost of not being able to separate its various components. The crucial missing link needed to answer this question is to what extent the intensive margin of college investment makes use of resources that are *rival* between students: if this not the case, most of the gains can be achieved without a large expansion of college resources. It is useful to discuss this distinction with two examples of intensive margins that have received attention in the literature: time use and quality of the college attended. If the most relevant frictions for low income students are those that prevent them from using their own time productively in college (for example because they work part time, they spend hours on the bus or simply they lack information on the most effective way to invest their time), then easing these barriers would allow to achieve gains without affecting the college opportunities of other students. On the contrast, since the capacity of high quality colleges is fixed (at least in the short run), empowering more students from poor background to attend places like Harvard or Princeton would require some of the current attenders to receive a lower quality education.[34]

Something that this framework can definitely document is that the elimination of barriers to college investment would not require a large expansion of the *number* of students attending college. As discussed in Section 2.5, the counterfactual increase in the college enrollment rate is minimal, and therefore it is reasonable to conjecture that most of the gains could be achieved even keeping this quantity fixed. However, further research is needed to clarify the extent to which the expansion of the intensive margin would require more resources to be invested in the US higher educational system.

[34]The model does not fully capture this dimension, since the supply of educational services is not explicitly characterized. However, decreasing returns in the production function endogenously put a boundary on how many educational units can be profitably accumulated.

2.7 Conclusions

Family income shapes the college careers of US students, even when its effect on pre-college human capital accumulation is accounted for. While this inequality of educational opportunities might be problematic for many reasons, in this paper I study its effect on a relatively under-appreciated dimension: productive efficiency.

I argue that a more meritocratic access to higher education might have large benefits in terms of aggregate output and average wages. In the baseline counterfactual experiment, I estimate a potential output gain of approximately 11%, and wage gains of approximately 12% and 9% for high school and college educated workers respectively. Most of these benefits can be achieved without a large expansion of the share of students attending college; instead, they mostly come from a more equal access to investment on the intensive margin, which includes elements like school quality, major and time use during college.

Which policies should be implemented in order to achieve all of this? This paper is rather silent on this dimension. An inherent cost of the reduced form approach adopted here is that it does not allow to separately identify the importance of different components of college investment, and therefore it cannot provide definite suggestions on the ones that should be targeted. Recent research shows that very simple and cost effective interventions that provide information to low income students can go a long way in reducing the disparity in educational outcomes (Hoxby and Turner, 2013); more micro studies are needed to verify whether other approaches might be equally or more effective.

This paper shows that the returns from finding and implementing the appropriate policies are potentially very high. In an era where the human capital boost due to the baby boom generation is fading out and new sources of growth are difficult to come by, a more equal access to educational resources might be exactly what is needed.

2.8 Tables and Figures

Table 2.1: Family Income and Type of Degree: Multinomial Logit

	Type of Degree			
	Associate	Bachelor	Graduate	Other
Log Family Income	0.121	0.464***	0.575***	0.034
	(0.147)	(0.136)	(0.207)	(0.337)
AQFT	0.012*	0.076***	0.114***	0.033**
	(0.006)	(0.006)	(0.011)	(0.013)

Notes: Table shows the coefficients from a weighted multinomial logit regression. Dependent variable indicates the type of degree obtained in college; omitted category is college dropouts. The sample is restricted to individuals with some college education. Additional controls include age, race, gender and urban status. Robust standard error in parentheses. Sample weights are provided by the NLSY79. $***$, $**$ and $*$ denote estimates significant at the 1%, 5% and 10% confidence level. Source: NLSY79

Table 2.2: Family Income and Type of Degree: Predicted Probabilities

	Type of Degree				
	Dropout	Associate	Bachelor	Graduate	Other
Low Income	0.336	0.186	0.345	0.084	0.046
Middle Income	0.277	0.168	0.410	0.107	0.039
High Income	0.238	0.153	0.450	0.125	0.033

Notes: Table shows the predicted probabilities for each type of degree from a multinomial logit model. Low, Middle and High Income refer to the first, second and third tercile of the family income distribution. Probabilities are evaluated at the average of Log Family Income within terciles, and at the overall sample average of the other regressors.

Table 2.3: College Shares by Group

Group	College Share Model	College Share Data
Low Income - Low Ability	0.27	0.28
Low Income - Middle Ability	0.44	0.48
Low Income - High Ability	0.75	0.77
Middle Income - Low Ability	0.34	0.30
Middle Income - Middle Ability	0.50	0.55
Middle Income - High Ability	0.73	0.83
High Income - Low Ability	0.43	0.32
High Income - Middle Ability	0.75	0.62
High Income - High Ability	0.85	0.88

Notes: College Share is the share of individuals with more than 12 years of education. Source: NLSY79.

Table 2.4: Relative Wages

	Relative Wage Model	Relative Wage Data
College / Non College Educated	1.53	1.48
High Ability / Low Ability	1.64	1.54
High Income / Low Income	1.53	1.52
College Educated		
High Ability / Low Ability	1.45	1.39
High Income / Low Income	1.25	1.32
Non College Educated		
High Ability / Low Ability	1.42	1.18
High Income / Low Income	1.23	1.21

Notes: College Educated are individuals with more than 12 years of education. High (low) ability and income refer to individuals in the third (first) tercile of the distribution of AFQT scores and family income. Source: NLSY79.

Table 2.5: Counterfactual Results

	Counterfactual Change
$\Delta Y/Y$ (%)	10.78
$\Delta w_H/w_H$ (%)	8.58
$\Delta w_L/w_L$ (%)	12.37
$\Delta P[S(i) = 1]$	2.25

Notes: Table shows the counterfactual changes in output, average wages in the low and high skill sector and college attendance rate.

Table 2.6: Decomposing the Wage Changes

	Low Skill Sector	High Skill Sector
Price (%)	10.12	-4.66
Quantity (%)	2.04	13.88
Wage (%)	12.37	8.58

Notes: Table shows the counterfactual changes in wages in the low and high skill sector decomposed between changes in quantity of efficiency units and changes in the price of an efficiency unit.

Table 2.7: Wage Changes by Group

	Wage Changes	
	Total (%)	Extensive Margin Only (%)
Low Income	20.89	3.68
Low Income - Low Ability	14.08	1.41
Low Income - Middle Ability	34.75	5.83
Low Income - High Ability	15.56	0.77
Middle Income	18.88	2.01
Middle Income - Low Ability	8.11	-0.70
Middle Income - Middle Ability	27.78	4.04
Middle Income - High Ability	19.28	1.33
High Income	-2.81	-3.13
High Income - Low Ability	1.18	-3.03
High Income - Middle Ability	-2.95	-2.49
High Income - High Ability	-3.84	-1.71

Notes: Table shows the counterfactual changes (in percentage terms) in average wages across income - ability groups. For the extensive margin results, only the group specific college attendance rate is allowed to adjust.

Table 2.8: Non-cognitive Skills by Group

	Rosenberg Scale	Rotter Scale
Low Income		
Low AFQT	59.97	55.01
Middle AFQT	64.69	61.12
High AFQT	66.55	63.54
Middle Income		
Low AFQT	60.64	57.18
Middle AFQT	62.08	59.85
High AFQT	66.73	63.85
High Income		
Low AFQT	62.97	57.71
Middle AFQT	65.77	62.21
High AFQT	67.53	65.82

Notes: Table shows the average scores in the Rosenberg Self-Esteem Scale and in the Rotter Locus of Control Scale for each income - cognitive ability group. Scores range from 0 to 100; the standard deviations are 16.01 (Rosenberg) and 17.15 (Rotter). Source: NLSY79.

Table 2.9: Counterfactual Results when Including Non-cognitive Ability

	Counterfactual Change
$\Delta Y/Y$ (%)	11.50
$\Delta w_H/w_H$ (%)	8.79
$\Delta w_L/w_L$ (%)	13.77
$\Delta P[S(i) = 1]$	2.60

Notes: Table shows the counterfactual changes in output, average wages in the low and high skill sector and college attendance rate when the measure of ability that includes non-cognitive skills is used.

Table 2.10: Counterfactual Results with α Estimated by OLS

	Counterfactual Changes	
	Baseline Ability	Cognitive & Non-cognitive Ability
$\Delta Y/Y$ (%)	10.69	11.53
$\Delta w_H/w_H$ (%)	8.53	8.63
$\Delta w_L/w_L$ (%)	12.28	14.13
$\Delta P[S(i) = 1]$	2.23	2.72

Notes: Table shows the counterfactual changes in output, average wages in the low and high skill sector and college attendance rate when α is calibrated as described in the text. Baseline Ability refers to the case where only AFQT scores are used; Cognitive & Non-cognitive Ability refers to the case where AFQT, Rosenberg and Rotten scores are used.

Table 2.11: Counterfactual Results with a Finer Classification of Ability

	Counterfactual Changes				
# of Ability Quantiles:	3	4	5	6	7
$\Delta Y/Y$ (%)	10.78	10.45	9.66	9.67	10.28
$\Delta w_H/w_H$ (%)	8.58	8.29	7.87	8.13	8.35
$\Delta w_L/w_L$ (%)	12.37	12.42	11.18	10.80	11.53
$\Delta P[S(i) = 1]$	2.25	2.36	2.07	1.92	2.04

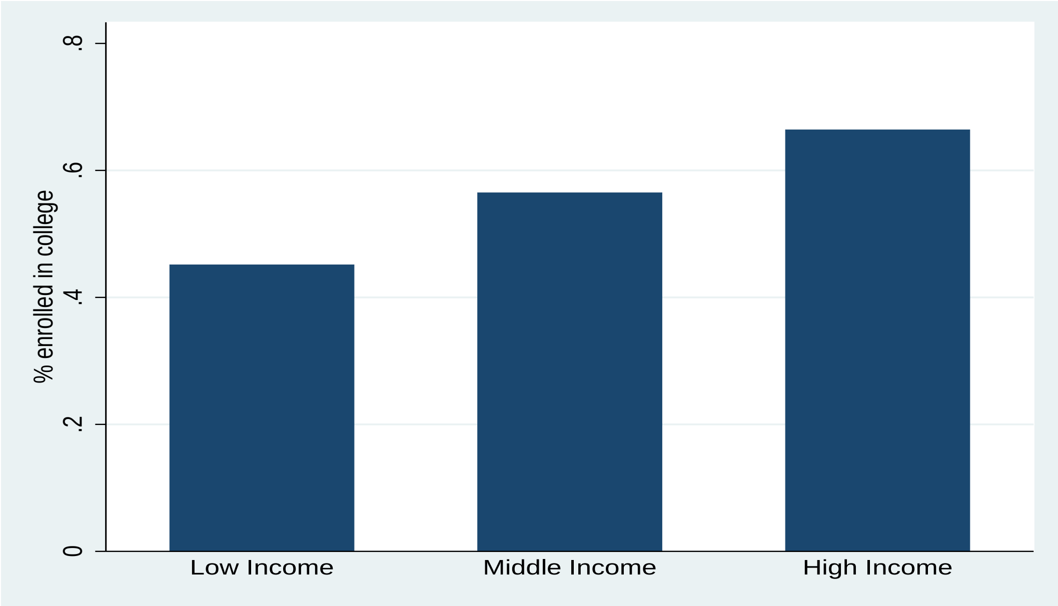
Notes: Table shows the counterfactual changes in output, average wages in the low and high skill sector and college attendance rate when the specified number of ability quantiles is used. Ability refers to the baseline case where only AFQT scores are used.

Table 2.12: Counterfactual Results with $\tau(y, z)$ Linear in Ability

	5^{th} percentile	95^{th} percentile
$\Delta Y/Y$ (%)	9.16	9.35
$\Delta w_H/w_H$ (%)	7.79	8.02
$\Delta w_L/w_L$ (%)	11.75	12.08
$\Delta P[S(i) = 1]$	1.87	1.98

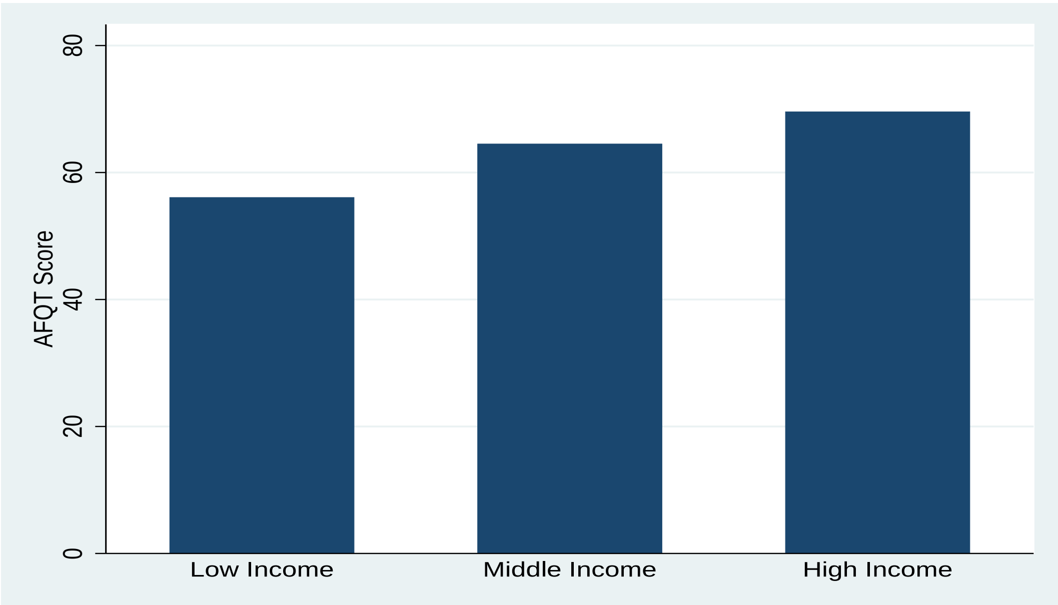
Notes: Table shows 5^{th} and 95^{th} percentiles of the simulated counterfactual changes in output, average wages in the low and high skill sector and college attendance rate. Ability refers to the baseline case where only AFQT scores are used.

Figure 2.1: College Enrollment by Family Income Terciles



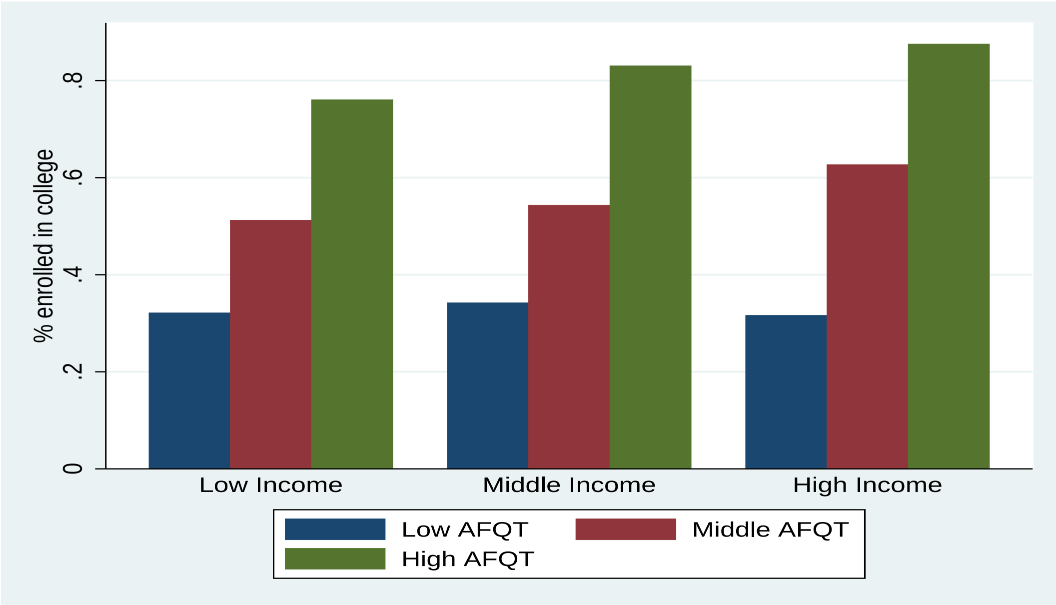
Notes: Height of the bar represents share with more than 12 years of schooling within each family income group. Source: NLSY79.

Figure 2.2: AFQT Scores by Family Income Terciles



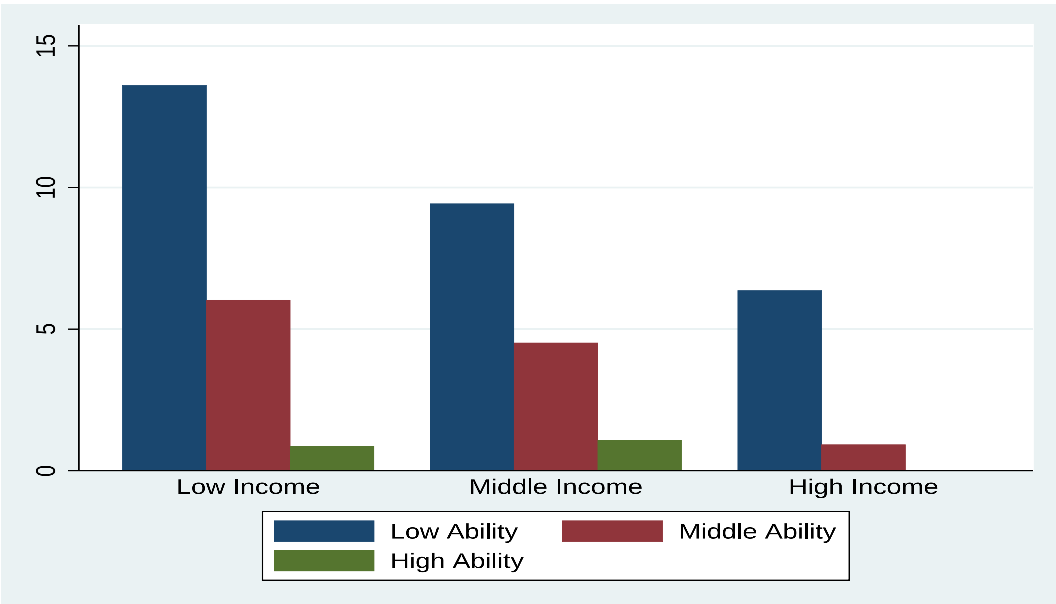
Notes: Height of the bar represents the average AFQT score within each family income group. Source: NLSY79.

Figure 2.3: College Enrollment by Family Income and AFQT Scores



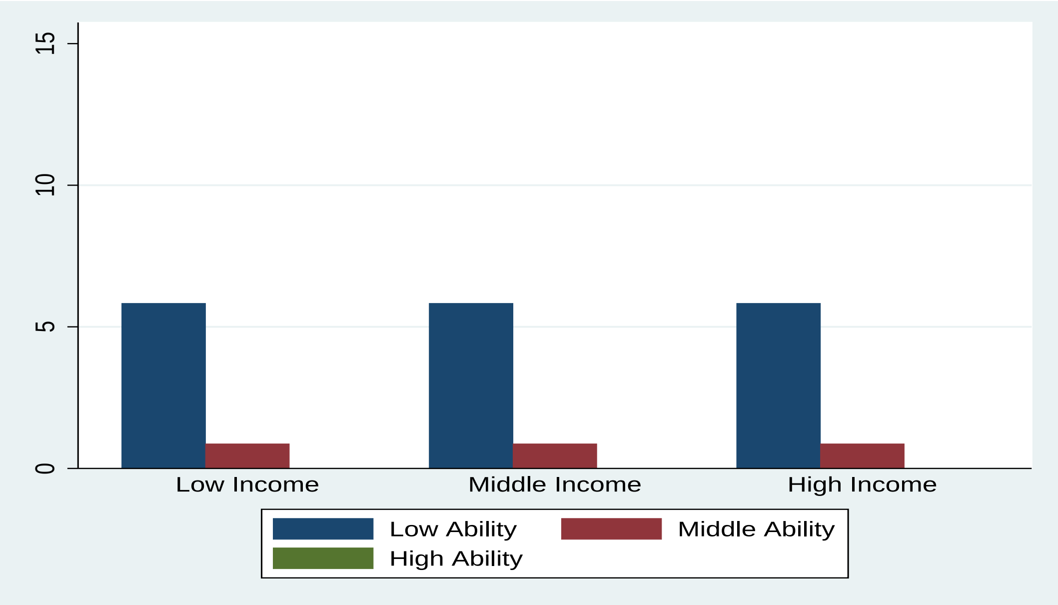
Notes: Height of the bar represents share with more than 12 years of schooling within each family income and test scores group. Source: NLSY79.

Figure 2.4: Estimated $\tau(y, z)$'s



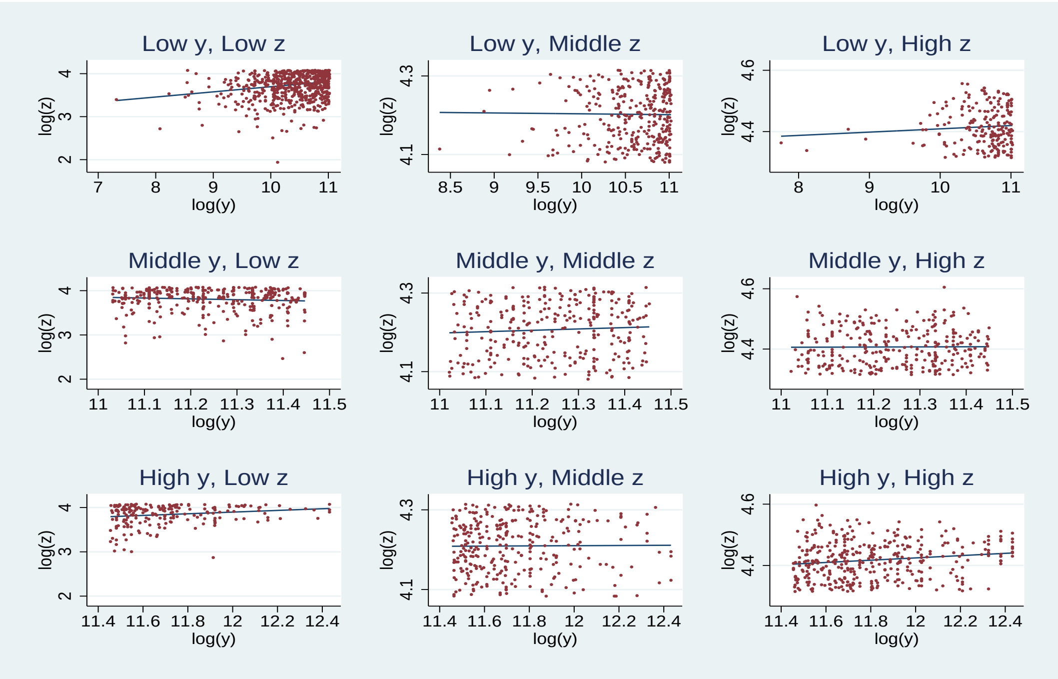
Notes: Height of the bar represents the estimated $\tau(y, z)$ for each income - ability group.

Figure 2.5: Counterfactual $\tau(y, z)$'s



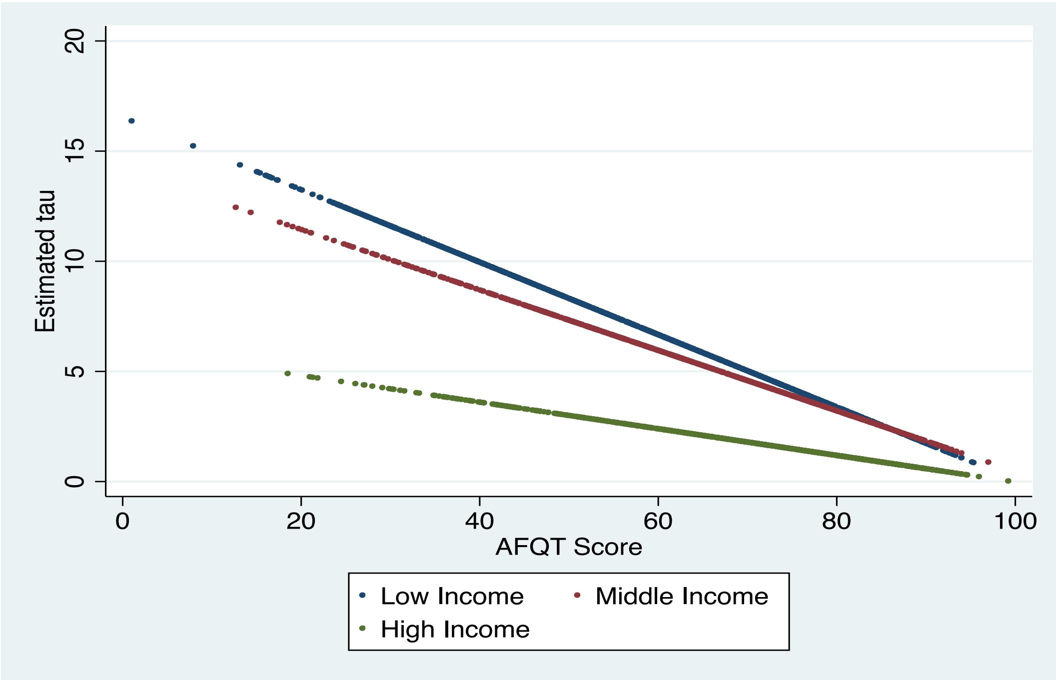
Notes: Height of the bar represents the counterfactual $\tau(y, z)$ for each income - ability group.

Figure 2.6: Ability and Family Income by Group



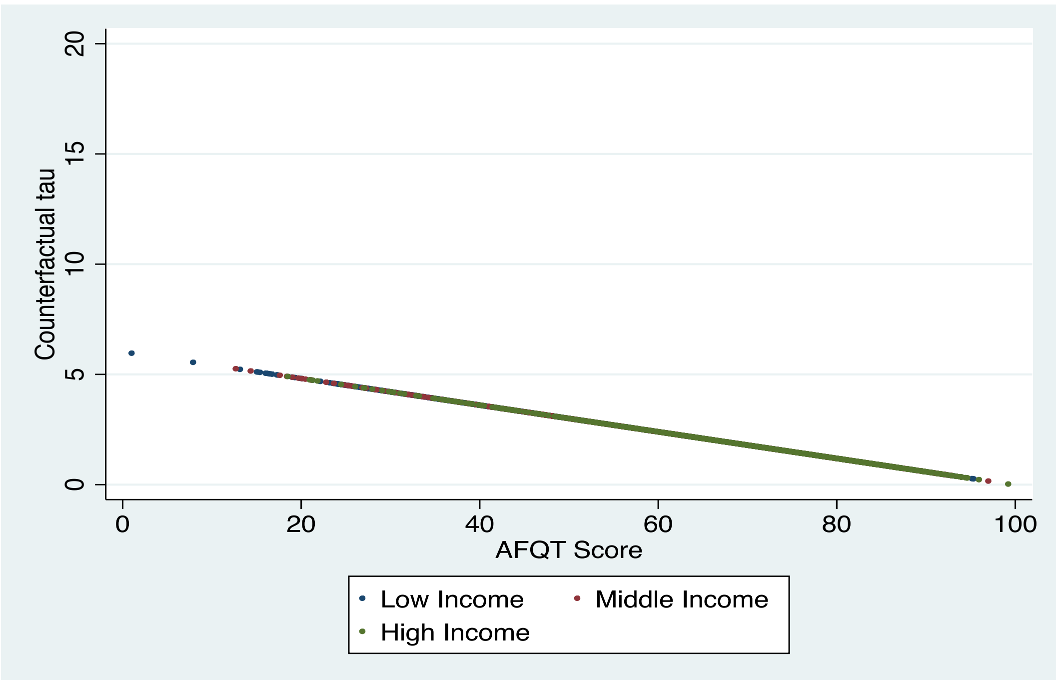
Notes: Each panel shows the (weighted) scatter plot of log ability (measured by AFQT scores) and log family income for the specified group. Line shows the best (weighted) linear fit.

Figure 2.7: Estimated $\tau(y, z)$'s - Linear in ability



Notes: Dots represent the estimated $\tau(y, z)$'s for each individual in the sample. Blue dots refer to the first family income terciles, red dots to the second and green dots to the third.

Figure 2.8: Counterfactual $\tau(y, z)$'s - Linear in ability



Notes: Dots represent the counterfactual $\tau^C(y, z)$'s for each individual in the sample. Blue dots refer to the first family income terciles, red dots to the second and green dots to the third.

Chapter 3

The Relative Efficiency of Skilled Labor across Countries: Measurement and Interpretation

3.1 Introduction

A question of major interest in macroeconomics is how the structure of production varies across countries. The traditional view is that rich and poor countries are set apart by large differences in a factor-neutral productivity shifter, while gaps in the relative amount and productivity of various factors of production are of more limited importance (Hall and Jones, 1999). Recently, this view has been challenged, thanks both to improved measurements of production inputs (Schoellman, 2012; Lagakos et al., 2016) and richer characterizations of the production technology (Jones, 2014; Caselli, 2016).

An emerging view in this line of research is that the relative efficiency of skilled and unskilled workers varies substantially across countries (Caselli and Coleman, 2006; Caselli, 2016; Malmberg, 2017). This conclusion typically follows from the analysis of quantities and prices. In a world with imperfect substitutability, a higher relative supply of skilled labor should be reflected in a lower relative price. However, existing estimates for the skill premium display limited variability across countries, in spite of large gaps in enrollment rates and educational achievements. This suggests that high-skilled workers are much more productive in rich (and skill-abundant) countries, attenuating the downward pressure on the skill premium stemming from their high supply. Cross-country gaps in the productivity of unskilled labor are instead moderate in size.

Different interpretations have been proposed to explain these patterns. One possibility, first advanced by Caselli and Coleman (2006) and Caselli (2016), is that technological differences across countries are factor-biased, and firms in rich countries adopt technologies more suitable for skilled workers. A natural alternative is that the human capital gap between high and low skilled workers is larger in rich countries, because of differences in educational quality, training or workers' intrinsic characteristics (Jones, 2014; Malmberg, 2017). In a cross-country setting, distinguishing between the two interpretations has important implications for various open questions in macro-development, such as the degree of transferability of technology across space and the role of human capital in accounting for cross-country gaps in economic performance.

In this paper I re-examine the measurement and intepretation of cross-country differences in relative skill efficiency. Using both aggregate and micro-level data, I confirm that gaps in the relative productivity of skilled and unskilled labor are large and not driven by the limited comparability or reliability of some of the sources used in previous studies. Building on this finding, I propose an approach based on the analysis of US immigrants to separately identify the role of technology and human capital in explaining the cross-country variation in relative skill efficiency.

The main data contribution of the paper consists in the construction of highly comparable estimates for the skill premium across countries. The lack of such information has represented a major drag on the existing literature, which has relied either on imputations based on related quantities, or on the use of sources not fully consistent with the underlying modelling strategy. To improve on this, I use micro-data from IPUMS International on 12 countries at different stages of development, ranging from the United States to India. I estimate the skill premium using the same specifications and similar sample restrictions for all countries. While the magnitude of some of the estimates is quite different from existing sources, I confirm the finding that the skill premium varies little across countries.

Through the lens of a simple production function setting, I back out the implied relative efficiency of skilled labor for each country, using both micro-data from IPUMS and more traditional sources to estimate the relevant parameters. I embed in this framework differences in both relative human capital and technology bias, and show that the estimated relative skill efficiency is a composite of the two. I confirm that relative skill efficiency varies substantially across countries. The productivity gap between skilled workers in the United States and in the average country of the sample is 4 times as large as the corresponding gap for unskilled workers.

I then move to the analysis of US immigrants. The basic idea of my approach is that comparing the within-group skill premia across immigrant groups, educated in their countries of origin but observed in the same labor market provides a way to isolate cross-country differences in relative human capital quality, keeping constant the local technological environment and other institutional characteristics. Gaps in the relative productivity of skilled labor might reflect differences in educational quality, as emphasized in Schoellman (2012), or differential of sorting into higher education across countries. I extend the baseline framework to incorporate immigrant workers with different levels of human capital, and then bring it to the data using the 2000 US Census. Conditional on the relevant controls, the difference between country of origin-specific skill premia identifies the corresponding gap in terms of relative skill quality.

I find that the cross-country variation in relative skill quality is of limited magnitude. While the productivity gap between skilled and unskilled workers is higher in the United States compared to most countries, the differences are much smaller than what it would be expected in a world where human capital quality explained the cross-country gaps in skill efficiency. Indeed, I conclude that differences in the skill bias of technology accounts for more than 90% of the cross-country variance in skill efficiency. While in principle patterns of differential selection into migration as a function of skills and country of origin might contribute to shape these results, I argue that this concern is unlikely to majorly affect the basic conclusion of the paper.

My work fits in the literature on cross-country differences in the structure of production. The basic approach to isolate skill-biased differences in productivity is introduced by Caselli and Coleman (2006), and subsequently updated by Caselli (2016). Recent work by Malmberg (2017) proposes an alternative methodology, based on trade data, to infer cross-country differences in the efficiency of skilled labor, and discusses the implications for development accounting. Compared to these papers, my main contributions are an improved measurement of skill premia and the development of a methodology to discriminate between relative skill quality and technology bias as sources of differences in skill efficiency. This distinction mirrors, on a cross-country dimension, a related debate on the relative roles of technology, human capital and sorting in explaining the rise of the skill premium over time (Acemoglu, 1998, 2002; Bowlus and Robinson, 2012; Hendricks and Schoellman, 2014).

This paper is also closely related to a growing literature studying the labor market experience of immigrants to learn about cross-country differences in human capital (Schoellman, 2012, 2016; Lagakos et al., 2016). In particular, Schoellman (2012) uses estimated Mincerian returns to schooling across immigrants' nationalities to quantify the role of educational quality for development accounting. While his focus is the aggregate human capital stock (in a model with perfect substitutability across skill levels), the main object of interest of my analysis is the relative quality of high skill and low skill workers. Immigrants from rich countries have higher returns both within and between skill levels, but the variation in returns between skill groups (which drive my estimates of relative skill quality) is more limited.

The paper is structured as follows. Section 3.2 introduces the basic framework and describes the measurement of relative skill efficiency. Section 3.3 shows evidence on immigrants, while Section 3.4 discusses the issue of selection. Finally, Section 3.5 concludes by discussing some implications and possibilities for future work.

3.2 Measuring Relative Skill Efficiency

In this section I document how the relative efficiency of skilled labor varies across countries. I introduce a simple framework, discuss how I bring it to the data and summarize the main patterns.

3.2.1 Framework

Suppose that country c is endowed with the following production technology,

$$Y_c = A_c K_c^{\alpha} \left[(A_{Hc} H_c)^{\rho} + (A_{Lc} L_c)^{\rho} \right]^{\frac{1-\alpha}{\rho}}$$

where K_c is physical capital and H_c and L_c are total high-skilled and low-skilled labor services. The production function involves three different technological paramenters, potentially varying across countries: A_c is total factor productivity, while A_{Hc} and A_{Lc} are factor biased technological terms, augmenting high- and low-skill labor. To simplify the notation, in what follows I omit the subscript c where this does not generate confusion.

The total amounts of high- and low-skill services used for production are

$$H = Q_H \tilde{H} \tag{3.1}$$

$$L = Q_L \tilde{L} \tag{3.2}$$

where $\tilde{H}$ and $\tilde{L}$ are the quantities of skilled and unskilled workers, while Q_H and Q_L represent their quality, or the amount of labor services provided by a given worker. While A_H and A_L proxy for factors external to individuals, such as the available technologies and the features of the working environment, I think of Q_H and Q_L as capturing workers' embodied productivity, which is possibly the result of both accumulated knowledge and innate characteristics.

Assuming perfectly competitive labor markets, the wage ratio between skilled and unskilled workers is

$$\frac{w_H}{w_L} = \left(\frac{A_H Q_H}{A_L Q_L}\right)^{\rho} \left(\frac{\tilde{H}}{\tilde{L}}\right)^{\rho-1} \tag{3.3}$$

I refer to $\frac{A_H Q_H}{A_L Q_L}$ as the relative "efficiency" of skilled and unskilled workers. If $\rho > 0$, which is the empirically relevant case given the existing estimates of the elasticity of substitution (Ciccone and Peri, 2005), a higher efficiency of skilled labor raises the skill premium, conditional on factor supplies. There are two reasons why this relative efficiency might vary across countries: differences in the skill bias of technology, $\frac{A_H}{A_L}$, and differences in the relative quality of skilled labor, $\frac{Q_H}{Q_L}$.

3.2.2 Measurement

In this section I describe how I map this framework to the data. The key object of interest is the relative efficiency of skilled labor $\frac{A_H Q_H}{A_L Q_L}$, which I normalize so that it is 1 for the United States. I take 2000 as my baseline year, and consider data sources relative to (or as close as possible to) this date.

A key choice to make is the split between high- and low-skill labor. Following most of the literature, I adopt a criterion based on workers' level of educational attainment. In particular, I consider skilled workers all high-school graduates and above, while individuals with at most some secondary schooling are unskilled. I chose this threshold mainly for two reasons: (i) comparability, as this is a definition of skilled labor commonly used in the literature on skill-bias technology differences (Caselli, 2016), and more broadly on human capital and economic development (Jones, 2014); (ii) the fact that, using the same split, Ciccone and Peri (2005) provide a credibly identified estimate for the elasticity of substitution between skilled and unskilled labor. Indeed, I use their estimated elasticity of 1.5 throughout, which implies a value for ρ of

1/3.

In practice, workers are obviously heterogeneous within these broad skill categories. I allow their productivity to depend on their level of educational attainment (indexed by j), gender (indexed by g) and experience (indexed by exp). For educational attainment, I split the unskilled in two groups (primary or less, some secondary) and the skilled in three groups (secondary completed, some tertiary, tertiary completed).[1] I define (potential) experience as the difference between current age and age at the end of education, and I consider nine groups based on 5-year intervals (0 to 4, 5 to 9, 10 to 14, 15 to 19, 20 to 24, 25 to 29, 30 to 34, 35 to 39, 40 or more). Within skill groups, human capital services provided by workers that differ in these dimensions are perfect subsitutes. I assume the aggregators $\tilde{H}$ and $\tilde{L}$ take the form[2]

$$\tilde{H} = \sum_{j \in H} \sum_{g} \sum_{exp} e^{\beta_j} e^{\lambda_g} e^{\mu_{exp}} n_{j,g,exp} \tag{3.4}$$

$$\tilde{L} = \sum_{j \in L} \sum_{g} \sum_{exp} e^{\beta_j} e^{\lambda_g} e^{\mu_{exp}} n_{j,g,exp} \tag{3.5}$$

where $n_{j,g,exp}$ is the number of workers belonging to the (j, g, exp) group, which I calculate from Barro and Lee (2013).[3] The parameters β_j, λ_g and $e^{\mu_{exp}}$ capture the productivity differentials (compared to a baseline category) associated to each educational achievement, gender and experience level. I normalize $\beta_{prim} = \beta_{sec} = \lambda_{male} = \mu_{0to4} = 0$, so that $\tilde{H}$ ($\tilde{L}$) is expressed in terms of equivalents of primary (secondary) educated, male and less than 5 years experienced workers (I refer to these groups as "baseline" skilled and unskilled workers). I start from a specification where β_j, λ_g and $e^{\mu_{exp}}$ do not vary across countries, but I also consider extensions where they do.

I then use the assumption of perfectly competitive labor markets to estimate β_j, λ_g and μ_{exp}. In particular, perfect competition implies that the average log wage of a worker of skill $S \in \{H, L\}$, with educational attainment j, gender g and experience exp is:

$$\log w_{S(j,g,exp)} = \alpha + \gamma_S + \beta_j + \lambda_g + \mu_{exp} \tag{3.6}$$

where α is a constant and $\gamma_S = \log (A_S Q_S)^{\rho} (\tilde{S})^{\rho - 1}$. The parameters β_j, λ_g and μ_{exp} can be therefore identified from a regression of log wages on skill group, educational attainment, gender and experience fixed effects.[4] Moreover, the coefficient on γ_H (with low-skilled workers

[1]While the cross-country data in Barro and Lee (2013) allows me to distinguish also between workers with no education, some primary and primary completed, I chose this broader aggregation for consistency with the other data sources of the paper. In particular, the US Census does not fully distinguish between some primary and primary completed. Moreover, sample sizes are small at these levels of educational attainment, especially by immigrants' countries of origin.

[2]With a slight abuse of notation, I denote by $j \in H$ ($j \in L$) the educational attainment levels assumed to be high- (low-) skill.

[3]In order to measure (potential) experience, I use data from the World Development Indicators on the country-specific statutory duration of each educational stage, as well as on the schooling starting age. For countries lacking this information, I use (World Bank defined) region-specific averages. Incomplete education spells are assigned half of the statutory duration, while for tertiary education (not covered in the WDI data) I use a duration of 4 years for all countries.

[4]Individual-level heterogeneity, resulting in an error in term in 3.6, can be easily added to the model; see

being the omitted cateogory) identifies the log skill premium, i.e. the log wage differential between baseline skilled and unskilled workers. I run this specification using data from the US Census, focusing on a sample of native individuals between 15 and 64 years old with a relatively high degree of labor market attachment.[5] Table 3.1 shows the estimated coefficients. Baseline skilled workers earn approximately 31% more than baseline unskilled workers. Within skill groups, wages rise steeply with educational attainment and experience, while females face a conditional wage gap of about 25%.

With the estimates of β_j, λ_g and μ_{exp} at hand I can compute $\tilde{H}$ and $\tilde{L}$ for all countries, so that the only missing part to back out $\frac{A_H Q_H}{A_L Q_L}$ from (3.3) is the skill premium $\frac{w_H}{w_L}$. To the best of my knowledge, no existing dataset provides a measure of the skill premium which is comparable across countries, nationally representative and consistent with the skill categorization used in this paper and the rest of the literature. Sources like ILOSTAT, compiled by the International Labor Organization, allow to contruct, for a limited number of countries, wage gaps between workers in different occupations or economic activities (as opposed to different educational attainments). This is problematic, as occupations and their skill content are difficult to compare across countries at different stages of development. Moreover, these data do not allow to condition in a systematic way on experience, age and other observable characteristics.

To overcome this problem, I adopt two complementary approaches. First, I follow much of the literature (Caselli and Coleman, 2006; Caselli, 2016) in inferring the skill premium from existing data on educational attainment and Mincerian returns to education. Second, I estimate directly the skill premium using micro-data for a smaller sample of countries.

For the first approach, I rely on a dataset I constructed in previous work with Francesco Caselli and Jacopo Ponticelli (Caselli et al., 2016), which includes a collection of estimated Mincerian returns for a large set of countries.[6] As shown in Caselli (2016), there is a direct mapping between the return estimated in a regression of low wages on years of schooling and the skill premium. In particular, if b is the Mincerian return, then

$$\log \frac{w_H}{w_L} = \frac{b \sum_j (s_j - \bar{s})^2 n_j + \sum_j \sum_g \sum_{exp} (\beta_j + \lambda_g + \mu_{exp})(s_j - \bar{s}) n_{j,g,exp}}{\sum_{j \in H} (s_j - \bar{s}) n_j} \tag{3.7}$$

where $n_j = \sum_g \sum_{exp} n_{j,g,exp}$ and s_j are the number and the completed years of schooling of workers with educational attainment j, while $\bar{s}$ is the average of attained years of schooling in the population. I compute $\log \frac{w_H}{w_L}$ from this expression using the estimates for β_j, λ_g and μ_{exp} illustrated above, data from Barro and Lee (2013) on educational attainment by gender and experience and from the World Bank's World Development Indicators on the duration of

Section 3.4. Of course, this specification might fail to capture causal effects, as several relevant unobservables are likely to be correlated with the regressors. The literature on returns to schooling, however, finds that OLS and IV estimates are often close in magnitude (Card, 2001).

[5] I restrict the sample to individuals that report working for wages, for at least 30 weeks and 30 hours per week in the previous year.

[6] The dataset contains Mincerian returns estimated on a large set of published and unpublished academic works. It contains up to two estimates for each country, one relative to the 1990s and one to the 2000s. Here I use the observation from the 2000s when possible, and the one from the 1990s otherwise.

each education stage (which I use to infer s_j and $\bar{s}$). This approach has obvious limitations, as it imposes assumptions on the human capital aggregators and it is based on estimates (the Mincerian coefficients) not in all cases coming from nationally representative samples. On the other hand, this methodology has the advantage of wide applicability: it allows me to construct the skill premium for 80 countries, spanning the whole income distribution. In what follows, I refer to this group of countries as the "broad" sample.

The second approach, which is a novel contribution of this paper, improves the reliability and comparability dimensions of the estimates at the cost of a smaller sample size. Using IPUMS, I estimate the skill premium using micro data for all countries with available information on wages or earnings, education, labor market status, gender and experience. This is possible for 12 countries in 2000 or a close year, including (according to the World Bank classification) high-income (United States, Canada, Israel, Trininidad and Tobago), upper middle-income (Mexico, Panama, Uruguay, Venezuela, Brazil, Jamaica) and lower middle-income (Indonesia, India) countries. I refer to this set of countries as the "narrow sample". For all of them, I run a specification equivalent to (3.6), imposing the same restriction as in the US sample whenever possible.[7] This also allows me to drop the assumption that β_j, λ_g and μ_{exp} do not vary across countries.[8]

3.2.3 Results

In this section I show how the relative efficiency of skilled labor varies across countries. I discuss and compare results on both the broad sample, where wage premia are inferred from Mincerian coefficients, and the narrow sample, where all quantities are directly estimated from micro-level data.

3.2.3.1 Broad Sample

Table 3.2 summarizes the variation of the skill premium, the relative supply of skilled and unskilled workers and the relative efficiency of skilled labor across the 80 countries in the broad sample. The wage ratio between skilled and unskilled workers varies substantially across countries, ranging from values close to or (in one case, Ukraine) below 1 to 11.6 in Jamaica. As expected, the correlation between the skill premium and relative skill supply (as well as GDP) is negative.

On the other hand, the relative supply of skilled labor is higher in rich countries. In the

[7]The information on weeks worked in the previous year is available only in the US, and therefore I cannot restrict the sample to workers with more than 30 weeks worked for the other countries. With the exception of India, Panama and Uruguay, which provide no information on labor supply, I can still restrict attention to individuals working more than 30 hours per week in the other countries. For Brazil, Mexico and Venezuela I use total earnings as wages are unavailable.

[8]In two countries, Israel and Jamaica, the information in IPUMS does not allow to identify individuals with some (not complete) secondary education, making it impossible to estimate the return to this level of educational attainment. As in Barro and Lee (2013)'s data the share of individuals with incomplete secondary education is positive, I impute their return interpolating the returns to primary and secondary education, using the returns to primary, some secondary and secondary education in the other 10 countries to construct the weights.

United States there are 12.71 as many baseline equivalent high-skilled workers as unskilled ones, while the corresponding number for Cambodia, the poorest country in the sample, is 0.04. This is partially because within the high-skill group US workers are relatively more concentrated in higher levels of educational attainment, blowing up the corresponding amount of baseline worker equivalents, and to a large extent because there are many more skilled workers in the US (the ratio between the numbers of skilled and unskilled workers, not converted to baseline equivalents, is 9.89 in the US and 0.03 in Cambodia).

The last row of Table 3.2 shows that there are large cross-country differences in the relative efficiency of skilled labor. The efficiency bias is 4 times larger in the US compared to the average country in the sample. Figure 3.1 displays in log scale the strong positive relationship between relative efficiencies and relative supplies: skilled labor is relatively more efficient where it is relatively more abundant.

The result is driven by the fact that the relationship between the skill premium and the relative supply of skilled workers is not steep enough, so that a high efficiency of skilled labor in skill abundant countries is needed to fit the data. Figure 3.2 illustrates this point by plotting the log skill premium against the log relative supply. The dashed line has a slope of $\rho - 1 = -0.67$, which is the predicted slope of this relationship in a world where $\log \frac{A_H Q_H}{A_L Q_L}$ was constant across countries (or, more generally, uncorrelated with $\log \frac{\tilde{H}}{\tilde{L}}$). The best linear fit (solid line) has instead an estimated slope of -0.17, with a standard error of 0.04. This implies that $\log \frac{A_H Q_H}{A_L Q_L}$ must increase with $\log \frac{\tilde{H}}{\tilde{L}}$.

3.2.3.2 Narrow Sample

Table 3.3 displays the skill premia, skill relative supplies and relative efficiencies for the countries in the narrow sample. As before, the skill premium is on average lower in countries with higher supply of skilled labor, but the range of its variation is relatively modest. Coupled with the large gaps in relative human capital displayed in the second column, this implies once again large cross-country differences in the relative efficiency of skilled labor. The magnitudes are similar to the ones for the broad sample: the average country has about one quarter of the skill efficiency bias compared to the US. Figure 3.3 shows that the relative efficiency skilled labor is strongly positively related to its relative supply.

Figure 3.4 illustrates the mechanics behind the result: the log skill premium is remarkably flat across countries (the slope is -0.10), in spite of the important variation in the log relative supply. To give an example, in a world where all countries had the US level of efficiency bias, the model would predict for Indonesia a wage ratio of 26, while the actual ratio is 2.12.

Taking stock, the analysis of micro data for a number of countries at different levels of development supports the existence of large gaps in skill efficiency, with richer and more skill abundant countries having relatively more efficient skilled labor. This pattern, both qualitatively and quantitatively, does not appear to be an artifact of differences in the data sources or in the measurement of the skill premium. This leads naturally to the next question: what explains the variation in relative skill efficiency across countries?

3.3 Sources of Differences in Relative Skill Efficiency

In this section I investigate the role of the skill bias of technology and the quality of skilled labor in explaining the cross-country variation in skill efficiency documented above. My strategy is based on the analysis of immigrants educated in different countries and observed in the same labor market. I first modify the baseline framework to include a specific role for workers' country of origin. I then map the new framework to the data and discuss the emerging patterns.

3.3.1 A Modified Framework

I introduce a new dimension of workers' heterogeneity to the framework in Section 3.2.1: the fact that some of them are educated in different countries. For clarity, I abstract from educational careers spanning more than one country, and I consider only natives and migrants entirely educated in their own country of origin.

I assume that skilled and unskilled workers' embodied human capital depends on the country where their education was acquired (indexed by a). This might reflect the combined impact of several characteristics of the educational environment, but also the mechanisms according to which individuals with different baseline characteristics sort into different levels of educational attainment. I do not wish (or need) to take a stand on the source of embodied productivity differences between skilled and unkilled labor, which might also be different across countries. I take as given their (possible) existence, and attempt to measure them in the data.

Within skill groups, services provided by different immigrant groups are perfect substitutes. The total quantities of high- and low-skill services used for production in country c are

$$H_c = \sum_a Q^a_{Hc} \tilde{H}^a_c \tag{3.8}$$

$$L_c = \sum_a Q^a_{Lc} \tilde{L}^a_c \tag{3.9}$$

where $\tilde{H}^a_c$ and $\tilde{L}^a_c$ are the number of (baseline equivalent) skilled and unskilled workers educated in country a and working in c, and Q^a_{Hc} and Q^a_{Lc} represent their average quality.

In a competitive labor market, the wage ratio between skilled and unskilled workers educated in a generic country b is

$$\frac{w^b_{Hc}}{w^b_{Lc}} = \left(\frac{A_{Hc}Q^b_{Hc}}{A_{Lc}Q^b_{Lc}}\right)^\rho \left(\frac{\sum_a (Q^a_{Hc}/Q^b_{Hc})\tilde{H}^a_c}{\sum_a (Q^a_{Lc}/Q^b_{Lc})\tilde{L}^a_c}\right)^{\rho-1} \tag{3.10}$$

This expression summarizes the key source of variation for my empirical strategy. Immigrant groups educated in their home countries face similar labor market conditions, both in terms of the degree of technological skill bias ($\frac{A_{Hc}}{A_{Lc}}$) and of the general equilibrium effect of aggregate skill supply (the second term of (3.10)), but are endowed with different Q's depending on their country of origin. Under some additional assumptions, by comparing skill premia across origin countries one can isolate cross-country differences in the relative quality of skilled and unskilled labor.

3.3.2 Measurement

In this section I describe how I map this framework to the data. The objective is to separately identify $\frac{A_{Hc}}{A_{Lc}}$ and $\frac{Q^c_{Hc}}{Q^c_{Lc}}$, in order to study the variability of both across countries. I normalize $\frac{A_{Hc}}{A_{Lc}}$ and $\frac{Q^c_{Hc}}{Q^c_{Lc}}$ so that they are 1 for the US.

I focus on the native and foreign-born workers living in the United States, observed in the 2000 Census. I once again restrict attention to workers between 15 and 64 years old, who have been working for wages for at least 30 weeks and 30 hours per week in the previous year. To isolate the role of education in the origin country, I only consider immigrants which are likely to have completed their education before relocating to the US: as in Schoellman (2012), I restrict the sample to those who migrated at least six years after the age at which they should have ended their studies, given their level of educational attainment. Moreover, when working with the broad sample, I focus my cross-country analysis on the 41 countries for which I observe at least 100 immigrants (satisfying the sample restrictions mentioned above) in each skill group, plus the United States.

As before, I consider human capital aggregators that take into account the heterogeneity in education, gender and experience,

$$\begin{aligned}\tilde{H}^a_c &= \sum_{j\in H}\sum_{g}\sum_{exp} e^{\beta_j} e^{\lambda_g} e^{\mu_{exp}} n^a_{c(j,g,exp)} \\ \tilde{L}^a_c &= \sum_{j\in L}\sum_{g}\sum_{exp} e^{\beta_j} e^{\lambda_g} e^{\mu_{exp}} n^a_{c(j,g,exp)}\end{aligned} \tag{3.11}$$

where $n^a_{c(j,g,exp)}$ is the number of workers in group (j, g, exp) educated in country a. The average log wage of a worker educated in a, of skill $S \in \{H, L\}$, with educational attainment j, gender g and experience exp is:

$$\log w^a_{Sc(j,g,exp)} = \alpha_c + \gamma_{Sc} + \log Q^a_{Sc} + \beta_j + \lambda_g + \mu_{exp} \tag{3.12}$$

where α_c is a constant and $\gamma_{Sc} = \log (A_{Sc})^\rho \left(\sum_a Q^a_{Sc}\tilde{S}_c\right)^{\rho-1}$. In a specification including skill group fixed effects, the interaction terms between skill group and country of origin fixed effects (with US natives as omitted category) identify $\log Q^a_{S,US} - \log Q^{US}_{S,US}$ for $S \in \{H, L\}$, from which $\log \frac{Q^a_{H,US}}{Q^a_{L,US}}$ can be calculated (recall that $\log \frac{Q^{US}_{H,US}}{Q^{US}_{L,US}}$ is normalized to 1). Moreover, β_j, λ_g and μ_{exp} are identified from the coefficients on educational attainment, gender and experience fixed effects.

For robustness, I also consider alternatives where the returns to (within skill groups) education and experience is country of origin-specific.[9] Given that sample sizes are small for

[9] Bratsberg (2002) and Schoellman (2012) document differences in country of origin-specific Mincerian returns for US immigrants, while Lagakos et al. (2016) argue for country-specific returns to experience. Note that the heterogeneity of the relative quality of skilled and unskilled labor already implies heterogeneous Mincerian returns. In future work, I plan to examine more systematically the extent to which heterogeneous returns to schooling are

some education-experience-country of origin cells, I use for this purpose a less flexible specification with linear (within skill groups) returns to years of schooling and quadratic returns to experience,

$$
\begin{aligned}
\tilde{H}_c^a &= \sum_{j \in H} \sum_{g} \sum_{exp} e^{\beta_H^a (s_j^a - s_{\text{sec}}^a)} e^{\lambda_g} e^{(\mu_1^a \eta_{exp} + \mu_2^a \eta_{\text{exp}}^2)} n_{c(j,g,exp)}^a \\
\tilde{L}_c^a &= \sum_{j \in L} \sum_{g} \sum_{exp} e^{\beta_L^a (s_j^a - s_{\text{pri}}^a)} e^{\lambda_g} e^{(\mu_1^a \eta_{exp} + \mu_2^a \eta_{exp}^2)} n_{c(j,g,exp)}^a
\end{aligned}
\tag{3.13}
$$

where β_H^a, β_L^a, μ_1^a, μ_2^a are the (country of origin-specific) returns to years of schooling (potentially different for skilled and unskilled workers), experience and experience squared, s_j^a is the number of years of schooling of workers with educational attainment j (achieved in country a) and η_{exp} is years of experience for group exp.[10] The units are still baseline category equivalents, and the baseline categories are males, primary educated or less (s_{pri}^a years of schooling) with no experience for unskilled workers and males, secondary educated (s_{sec}^a years of schooling) with no experience for skilled workers. Similarly to (3.12), a regression of log wages on skill group fixed effects, years of schooling, experience, experience squared (all interacted by country of origin fixed effects) and gender identifies $\log \frac{Q_{Hc}^a}{Q_{Lc}^a}$, β_H^a, β_L^a, μ_1^a, μ_2^a for all countries of origin. In what follows, I refer to (3.13) and close variations as the "parametric" specifications.

Under the assumption that the relative quality of skilled workers among US immigrants captures the relative quality among natives in the country origin, that is $\log \frac{Q_{H,US}^a}{Q_{L,US}^a} = \log \frac{Q_{Ha}^a}{Q_{La}^a}$, I can examine the cross-country variation in the latter. The main question of interest is the role of relative skill quality in explaining differences in relative skill efficiency. Given that workers' quality is assumed to be heterogeneous depending of the country in which they were educated, in principle one should take into account the educational composition of the population in each country when computing relative skill quantities and backing out relative efficiencies. However, if immigrants educated abroad are a sufficiently small share of the working population, the relative supply, quality and price of skills among native workers are good approximations for the corresponding population-wide quantities.[11] I rely on this approximation and compute for each country $\frac{A_{Hc}}{A_{Lc}}$ from (3.3), using estimates for the relative skill quality among native workers.[12,13]

driven by differences within as opposed to between skill groups.

[10]I experimented with higher degree polynomials obtaining very similar results.

[11]More precisely, the population-wide skill premium is given by

$$
\frac{w_{Hc}}{w_{Lc}} = \left(\frac{A_{Hc} Q_{Hc}}{A_{Lc} Q_{Lc}} \right)^{\rho} \left(\frac{\sum_a (Q_{Hc}^a / Q_{Hc}) \tilde{H}_c^a}{\sum_a (Q_{Lc}^a / Q_{Lc}) \tilde{L}_c^a} \right)^{\rho - 1}
$$

where, for $S \in \{H, L\}$, $w_{Sc} = \sum_a w_{Sc}^a \frac{n_{Sc}^a}{n_{Sc}}$, $Q_{Sc} = \sum_a Q_{Sc}^a \frac{n_{Sc}^a}{n_{Sc}}$ and n_{Sc} (n_{Sc}^a) is the number of workers of skill S in the population (educated in country a). If $n_{Sc}^c \approx n_{Sc}$ for $S \in \{H, L\}$, then clearly $\frac{w_{Hc}}{w_{Lc}} \approx \frac{w_{Hc}^c}{w_{Lc}^c}$, $\frac{Q_{Hc}}{Q_{Lc}} \approx \frac{Q_{Hc}^c}{Q_{Lc}^c}$ and $\frac{\sum_a (Q_{Hc}^a / Q_{Hc}) \tilde{H}_c^a}{\sum_a (Q_{Lc}^a / Q_{Lc}) \tilde{L}_c^a} \approx \frac{\tilde{H}_c^c}{\tilde{L}_c^c}$.

[12]The Barro and Lee (2013)'s data, used to compute human capital stocks, refer to the whole population (natives and immigrants). The skill premium estimated from IPUMS data (for the countries in the narrow sample) is relative to native workers only (though including immigrants has a negligible impact on the resulting estimates).

[13]In principle, using data on the stock of migrants by country of origin, one could make some progress towards quantifying the importance of differences in the ethnic composition of the population. This approach would require,

From now on, I simply refer to these objects as $\frac{A_H}{A_L}$ and $\frac{Q_H}{Q_L}$.

To summarize the empirical strategy, I start from a difference-in-differences approach, where I compare, within the United States, the log wages of skilled and unskilled workers between the different countries where they were educated. I then examine whether skill premia are larger for countries of origin with a higher measured relative efficiency of skilled labor, and draw the implications for the cross-country dispersion in the latter.

3.3.3 Results

In this section I show how the relative skill bias of technology and quality of skilled labor vary across countries. I start from results relative to the broad sample and then consider the narrow one.

3.3.3.1 Broad Sample

Table 3.4 provides summary statistics on the relative skill efficiency, skill bias of technology and skill quality in the countries with at least 100 immigrant workers per skill group in the regression sample (and the United States). The first row shows that these countries display similar patterns in terms of skill efficiency as the rest of the broad sample (compare with the third row of Table 3.2). The second and third columns show that the skill bias of technology is substantially more dispersed across countries compared to relative skill quality. In the average country, the skill bias of technology is 31% of the US level, while relative skill quality is 88%. In Cambodia, a country with 27 times as many unskilled as skilled workers, the relative skill quality is 83% compared to the US.

Figure 3.5 illustrates how the skill bias of technology and relative skill quality vary as a function of skill supply (on a log scale). Countries with high relative supply of skilled workers have technologies more biased towards them (left panel), while the relationship with relative skill quality is almost flat (right panel).

As an alternative way to summarize these patterns, I consider the relative role of the two components of (log) skill efficiency in explaining its variance across countries and covariance with other variables of interest. In particular, I compute

$$V_A = \frac{V(\log \frac{A_H}{A_L})}{V(\log \frac{A_H Q_H}{A_L Q_L})} \quad , \quad V_Q = \frac{V(\log \frac{Q_H}{Q_L})}{V(\log \frac{A_H Q_H}{A_L Q_L})}$$

which give me the shares of the variance of log relative skill efficiency accounted for by tech-

across different host countries, information on the composition by education and age of arrival of the stock of migrants from each country of origin, and assumptions on the quality of individuals whose educational career spans more than one country. Given the substantial data requirements, the additional structure that this would involve and the fact the immigrants educated abroad are a small share of the population in most countries, I chose not to follow this route.

nology skill bias and relative skill quality, and

$$Cov_A\left(\frac{\tilde{H}}{\tilde{L}}\right) = \frac{Cov(\log\frac{A_H}{A_L}, \log\frac{\tilde{H}}{\tilde{L}})}{Cov(\log\frac{A_H Q_H}{A_L Q_L}, \log\frac{\tilde{H}}{\tilde{L}})} \quad , \quad Cov_A\left(y\right) = \frac{Cov(\log\frac{A_H}{A_L}, \log y)}{Cov(\log\frac{A_H Q_H}{A_L Q_L}, \log y)}$$

$$Cov_Q\left(\frac{\tilde{H}}{\tilde{L}}\right) = \frac{Cov(\log\frac{Q_H}{Q_L}, \log\frac{\tilde{H}}{\tilde{L}})}{Cov(\log\frac{A_H Q_H}{A_L Q_L}, \log\frac{\tilde{H}}{\tilde{L}})} \quad , \quad Cov_Q\left(y\right) = \frac{Cov(\log\frac{Q_H}{Q_L}, \log y)}{Cov(\log\frac{A_H Q_H}{A_L Q_L}, \log y)}$$

which represent the shares of the covariance between log relative skill efficiency and log relative skill supply on one hand and log GDP per worker on the other driven by $\log\frac{A_H}{A_L}$ and $\log\frac{Q_H}{Q_L}$.

Table 3.5 shows the results. Starting from the baseline specification (first column), the technology term accounts for the vast majority of both the variance of log skill efficiency (92%) and of its covariance with log skill supply (95%) and log GDP (81%). The covariance between $\log\frac{A_H}{A_L}$ and $\log\frac{Q_H}{Q_L}$ is small and positive (since V_A and V_Q almost sum to 1). The second column considers the parametric specification of human capital stocks, with linear returns to schooling and quadratic to experience; reassuringly, this modification does not affect the results. Then, the third column uses the parametric specification with country-specific returns to education and experience, as in (3.13).[14] The contribution of relative skill quality increases slightly, but overall the technology term still dominates.[15]

3.3.3.2 Narrow Sample

I now turn to the corresponding results for the narrow sample. Table 3.6 shows the patterns for skill efficiency, technology bias and skill quality. As for the broad sample, in terms of cross-country variation the technology term dwarfs the quality one. The average country displays 30% of the technology bias of the United States, and 92% of its relative skill quality. Figure 3.6 shows that more generally the relative skill bias of technology is strongly increasing in the relative supply of skilled workers, while the relative quality term is not.

An example is instructive to understand how the magnitude of the skill premium within immigrant groups is driving the result. Mexico's relative skill efficiency is 17% of the US level. If this gap was entirely reflecting higher relative skill quality among US workers, the skill premium among Mexican educated workers in the US should be 0.77 log points smaller than the one among US natives (corresponding to a wage ratio of 63%), while in reality the difference is 0.19 log points (wage ratio of 112%).

3.4 Selection

Given that my strategy consists in using immigrant workers to estimate country-specific differences in the relative quality of skilled labor, a natural concern is that emigrant workers are not

[14] Here, when computing the log skill premium I adjust the epression in (3.7) to be consistent with the different formulation of human capital stocks. The impact of this adjustment is negligible.

[15] While, within skill groups, returns to schooling (and, to some extent) experience are generally higher in richer countries, this raises both $\tilde{H}$ and $\tilde{L}$ having a limited impact on the relative skill supply.

randomly selected. In this section I discuss the possible consequences of selection and discuss some evidence on its importance.

It is helpful to explicitly introduce some individual-level heterogeneity in the framework of section 3.3.1 to illustrate the main issues. Suppose that the quality of individual i, of skill $S \in \{H, L\}$, having completed his education in country a is $Q^a_S \varepsilon^a_{S,i}$, where Q^a_S is a term common to all individuals of skill S educated in a and $\varepsilon^a_{S,i}$ captures the heterogeneity in unobservable skills. For analytical convenience, I assume that $\varepsilon^a_{S,i}$ follows a log-normal distribution with log-mean 0 and log-variance $(\sigma^a)^2$. Moreover, I mantain the assumption that $\varepsilon^a_{S,i}$ is uncorrelated with workers' observable characteristics (education, gender and experience).[16]

If migrants are selected on unobservable skills, $\mathbb{E}\left[\log \varepsilon^a_{S,i} | \text{migrant}\right] \neq 0$. The relative log skill quality I estimate out of US migrants using (3.11) would then read

$$\log Q^a_{H,US} - \log Q^a_{L,US} = \log Q^a_H - \log Q^a_L + \mathbb{E}\left[\log \varepsilon^a_{H,i} | \text{migrant}\right] - \mathbb{E}\left[\log \varepsilon^a_{L,i} | \text{migrant}\right]$$

which differs from the quantity of interest as long as $\mathbb{E}\left[\log \varepsilon^a_{H,i} | \text{migrant}\right] \neq \mathbb{E}\left[\log \varepsilon^a_{L,i} | \text{migrant}\right]$. Migrants' selection is therefore problematic to the extent that it takes place with a different degree across skill groups.

To the best of my knowledge, the migration literature offers little direct guidance on whether this form of differential selection might be relevant in practice. It has been widely established that migrants are non-randomly selected on observable and unobservable skills (Borjas, 1987), and for the vast majority of origin countries the degree of selection of emigrants to the United States appears to be positive (Feliciano, 2005b). However, less is known about the relative selection by educational achievement, i.e. on how, among individuals educated in a given country, the degree of selection on unobservables of migrants within the low-skill group compares to the one within the high-skill group.

Since my main result is that, for most countries, the log relative quality of skilled labor inferred out of migrants is too large to account for the international gaps in skill efficiency, investigating the possibility that $\mathbb{E}\left[\log \varepsilon^a_{H,i} | \text{migrant}\right] > \mathbb{E}\left[\log \varepsilon^a_{L,i} | \text{migrant}\right]$ is of particular interest. A more positive degree of selection across skilled workers could in principle lead me to understate the importance of relative skill quality differences across countries.

I discuss one piece of evidence suggesting that this is unlikely to be a major concern: the propensity to migrate conditional on skill group is much higher for the high-skilled than for the low-skilled. Figure 3.7 plots the share of skilled workers among US emigrants for each country of origin against the share of skilled workers in the country of origin population. Almost all countries are above the 45 degree line, showing that emigrants are substantially more likely to be high-skilled.[17] If, as suggested by the literature, migrants' selection on unobservables (conditional on observables) is positive, such a pattern would imply that skilled migrants are relatively more negatively selected. To see why, suppose there exists a skill-specific threshold

[16] This is obviously a strong assumption, though common in the development accounting literature. See footnote 4 for a related discussion.

[17] Similar patterns hold when expressing units in terms of baseline equivalent workers as opposed to counting persons.

t_S^a such that workers migrate if $Q_S^a \varepsilon_{S,i}^a \geqslant t_S^a$. The within-skill group share of emigrants is then $1 - \Phi\left(\frac{\log t_S^a - \log Q_S^a}{\sigma^a}\right)$, where $\Phi(.)$ is the standard Normal's cumulative distribution function. The fact that $1 - \Phi\left(\frac{\log t_H^a - \log Q_H^a}{\sigma^a}\right) > 1 - \Phi\left(\frac{\log t_L^a - \log Q_L^a}{\sigma^a}\right)$, implied by Figure 3.7, means that $\log t_H^a - \log Q_H^a < \log t_L^a - \log Q_L^a$. It follows that

$$\mathbb{E}\left[\log \varepsilon_{H,i}^a | \text{migrant}\right] = \frac{\sigma^a \phi\left(\frac{\log t_H^a - \log Q_H^a}{\sigma^a}\right)}{1 - \Phi\left(\frac{\log t_H^a - \log Q_H^a}{\sigma^a}\right)} < \mathbb{E}\left[\log \varepsilon_{L,i}^a | \text{migrant}\right] = \frac{\sigma^a \phi\left(\frac{\log t_L^a - \log Q_L^a}{\sigma^a}\right)}{1 - \Phi\left(\frac{\log t_L^a - \log Q_L^a}{\sigma^a}\right)}$$

Intuitively, if migrating is relatively easier for high-skilled individuals, the low-skilled ones who do migrate should be comparatively better selected on unobservable skills.

I conclude this section by noticing that the threat of selection is less severe when comparing countries other than the United States. This is because, while US estimates are based on native workers, for all other countries the relative skill quality is inferred from immigrants only. Cross-country gaps in these estimates capture actual gaps in relative skill quality as long as there is not a pattern of *differential relative selection* of skilled and unskilled migrants across countries, while a common degree of relative selection in unconsequential. As it is evident from Figures 3.5 and Figure 3.6, the result that technology skill bias (as opposed to relative skill quality) is the key factor varying between skill-poor and skill-abundant countries is not driven by the United States only.

3.5 Conclusions

In this paper I re-visit the question of how the relative productivity of skilled and unskilled labor differs across countries. I show that, according to various sources, the skill premium varies little across countries, implying large gaps in relative skill efficiency. In the second part of the paper, I show that skill premia within immigrant groups are not consistent with the view that differences in relative skill quality play a quantitatively important role. The variation in relative skill efficiency is instead more likely to be related to technological factors.

These results have important implications when considering the relative role of human capital and technology in accounting for cross-country differences in output per worker. Malmberg (2017) suggests that gaps in skill efficiency are an important component of differences in economic performance. My findings imply that one should be careful in attributing these gains to human capital. Moreover, if we accept the view that the factor-bias of adopted technologies is very different between rich and poor country, this gives credit to the possibility that rich countries' technologies might not be appropriate for firms in poor countries, as argued by Acemoglu and Zilibotti (2001).

Indeed, my results emphasize the importance of understanding the determinants of technological skill bias. A common view is that differences in the technology mix reflect the optimal responses of firms to the abundance or scarcity of skilled labor (Caselli, 2016). It would be useful to have a sense of the quantitative importance of this mechanism, and of whether other institutional, cultural or geographical factors might contribute to explain why poorer countries adopt less skill-biased technologies.

The approach of this paper can be extended in various directions. For example, it would be interesting to explore the relative role of technology and human capital in explaining the differential evolution of the skill premium over time in the United States and Europe. Moreover, a similar exercize could be performed within countries, in order to explore how relative skill quality vary across regions with different characteristics. I hope to address some of these open issues in future work.

3.6 Tables and Figures

Table 3.1: Returns to Skill, Education, Gender and Experience - US Census

Skill		Edu		Gender		Experience	
Group	Coeff	Group	Coeff	Group	Coeff	Group	Coeff
Low Skill	0	Primary or less	0	Male	0	0 to 4	0
High Skill	0.305 (0.005)	Some Secondary	0.111 (0.005)	Female	-0.251 (0.001)	5 to 9	0.289 (0.001)
		Secondary	0			10 to 14	0.451 (0.001)
		Some Tertiary	0.178 (0.001)			15 to 19	0.540 (0.001)
		Tertiary	0.615 (0.001)			20 to24	0.586 (0.001)
						25 to 29	0.610 (0.001)
						30 to 34	0.639 (0.001)
						35 to 39	0.648 (0.001)
						40+	0.636 (0.002)

Notes: The Table shows the estimated returns to skill, education, gender and experience in the 2000 US Census. The dependent variable is log wage per hour. The sample includes 4282320 observations. The coefficients on low skill, primary education or less, secondary education, male and 0 to 4 experience are normalised to 0. Observations are weighted according to the provided sample weights. Robust standard errors are shown in parentheses.

Table 3.2: Skill Premium, Supply and Efficiency across Countries - Broad Sample

	N	Mean	Min	Max	Std. Dev.	Corr w/ $\tilde{H}/\tilde{L}$	Corr w/ y
w_H/w_L	80	2.30	0.81	11.55	1.43	-0.27	-0.35
$\tilde{H}/\tilde{L}$	80	1.36	0.02	12.71	1.80	1	0.40
$(A_H Q_H)/(A_L Q_L)$	80	0.26	0.02	1.25	0.20	0.63	0.26

Notes: The Table shows summary statistics for the skill premium, relative skill supply and efficiency across the countries in the broad sample. Relative skill efficiency is normalised such that it takes value 1 for the United States.

Table 3.3: Skill Premium, Supply and Efficiency across Countries - Narrow Sample

Country	w_H/w_L	$\tilde{H}/\tilde{L}$	$(A_H Q_H)/$ $(A_L Q_L)$
Indonesia	2.12	0.15	0.08
Brazil	2.22	0.33	0.14
Venezuela	1.24	0.38	0.09
Uruguay	1.81	0.41	0.14
India	2.57	0.42	0.19
Mexico	2.09	0.45	0.17
Trinidad and Tobago	1.68	0.48	0.14
Jamaica	1.46	0.82	0.17
Panama	2.06	0.94	0.27
Canada	1.57	1.73	0.30
Israel	1.62	2.32	0.38
United States	1.36	12.71	1
Average	1.82	1.76	0.26

Notes: The Table shows the the skill premium, relative skill supply and efficiency across the countries in the narrow sample. Relative skill efficiency is normalised such that it takes value 1 for the United States.

Table 3.4: Relative Technology and Skill Quality across Countries - Broad Sample

	N	Mean	Min	Max	Std. Dev.	Corr w/ $\tilde{H}/\tilde{L}$	Corr w/ y
$(A_H Q_H)/(A_L Q_L)$	42	0.28	0.04	1.25	0.22	0.58	0.17
A_H/A_L	42	0.31	0.05	1.30	0.22	0.57	0.10
Q_H/Q_L	42	0.88	0.72	1.27	0.11	0.15	0.43

Notes: The Table shows summary statistics for relative skill efficiency, the relative skill bias of technology and relative skill supply across the countries in the broad sample. Only countries with at least 100 unskilled and 100 skilled workers in the regression sample are included. All variables are normalised such that they take value 1 for the United States.

Table 3.5: Variance and Covariance Decomposition

	Baseline	Parametric	Parametric + Country-Specific Returns
V_A	0.92	0.91	0.87
V_Q	0.04	0.04	0.06
$Cov_A\left(\frac{\tilde{H}}{\tilde{L}}\right)$	0.95	0.94	0.89
$Cov_Q\left(\frac{\tilde{H}}{\tilde{L}}\right)$	0.05	0.06	0.11
$Cov_A\left(y\right)$	0.81	0.79	0.74
$Cov_Q\left(y\right)$	0.19	0.21	0.26
N	42	42	42

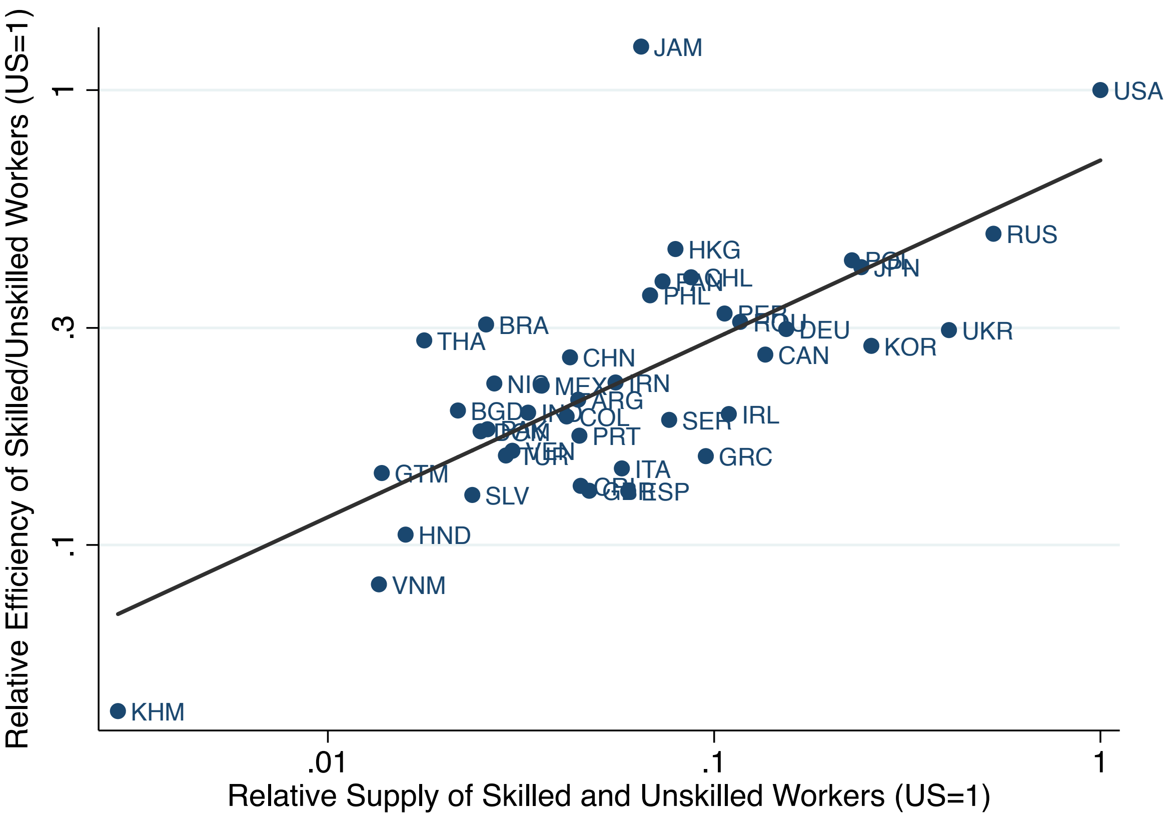
Notes: The Table shows results for the decompositions of the skill efficiency variance (first panel) and its covariance with log relative skill supply (second panel) and with log GDP (second panel). Only countries with at least 100 unskilled and 100 skilled workers in the regression sample are included. The second column refers to the specification with linear return to years of schooling and quadratic to experience, while the third refers to the specification where returns are country-specific.

Table 3.6: Relative Technology and Skill Quality across Countries - Narrow Sample

Country	$\frac{A_H Q_H}{A_L Q_L}$	$\frac{A_H}{A_L}$	$\frac{Q_H}{Q_L}$
Indonesia	0.08	0.10	0.85
Brazil	0.14	0.16	0.88
Venezuela	0.09	0.11	0.79
Uruguay	0.14	0.15	0.89
India	0.19	0.21	0.95
Mexico	0.17	0.20	0.83
Trinidad and Tobago	0.14	0.16	0.88
Jamaica	0.17	0.18	0.96
Panama	0.27	0.30	0.88
Canada	0.30	0.31	0.97
Israel	0.38	0.34	1.13
United States	1	1	1
Average	0.26	0.27	0.92

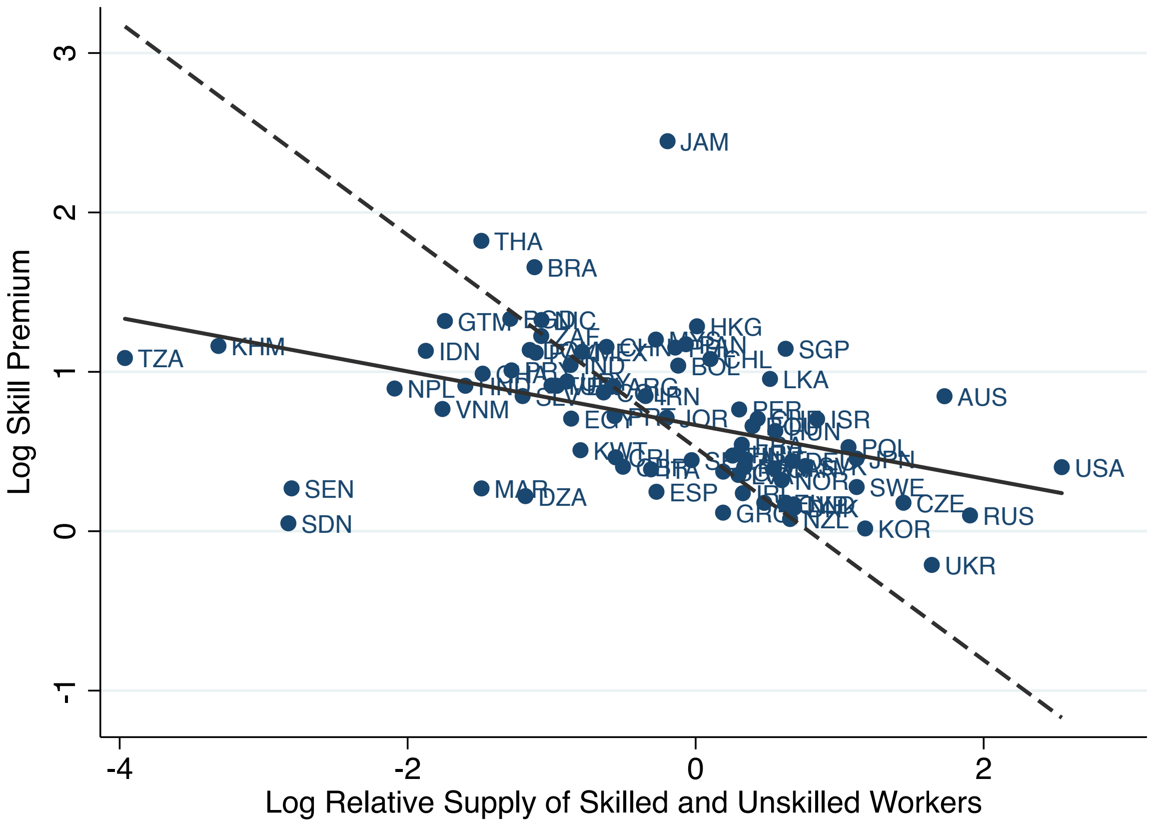
Notes: The Table shows summary statistics for relative skill efficiency, the relative skill bias of technology and relative skill supply across the countries in the narrow sample. All variables are normalised such that they take value 1 for the United States.

Figure 3.1: Relative Efficiency and Relative Supply of Skilled Labor - Broad Sample



Notes: The figure plots on a log scale the relative efficiency and relative supply of skilled workers for countries in the broad sample. Both variables are normalized so that they take value 1 for the United States. The solid line represents the best exponential fit.

Figure 3.2: Skill Premium and Skill Supply - Broad Sample

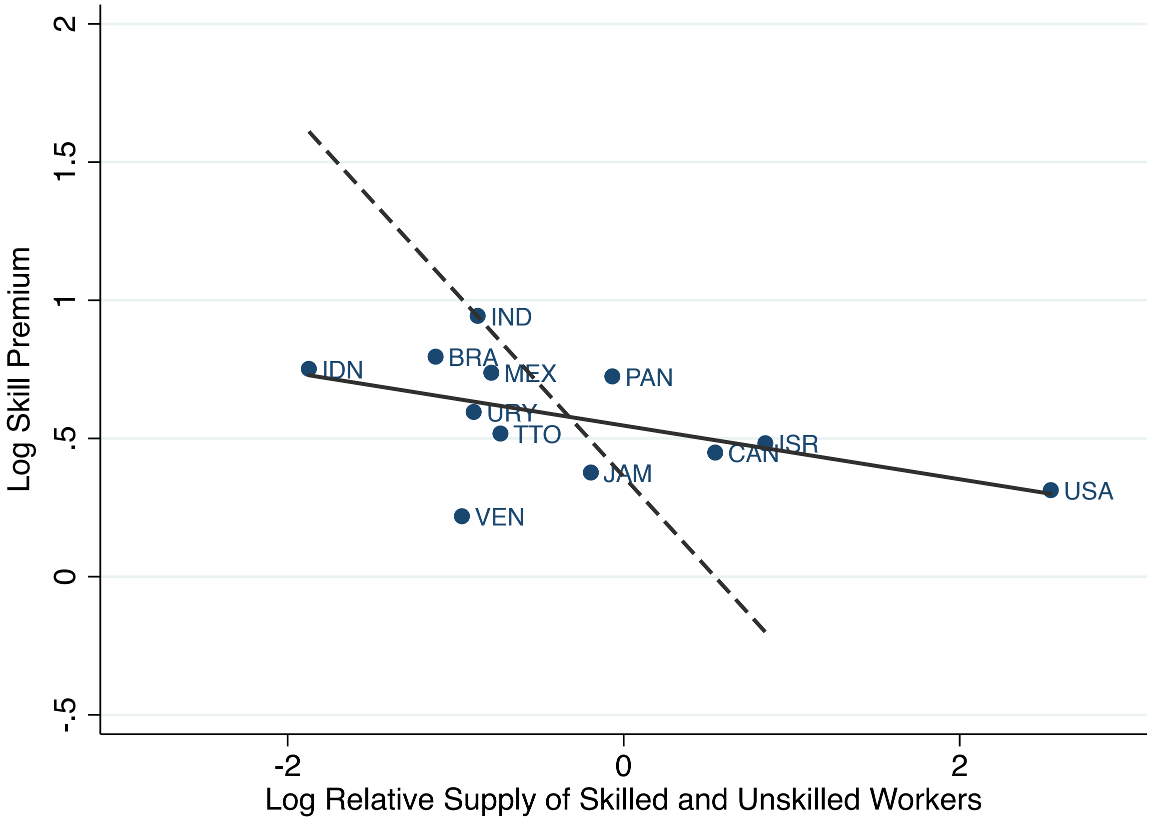


Notes: The figure plots the log skill premium and the log and relative supply of skilled workers for countries in the broad sample. The solid line represents the best linear fit. The dashed line has the slope of the predicted relationship (-0.67) in a counterfactual where skill efficiency and supply are uncorrelated.

Figure 3.3: Relative Efficiency and Relative Supply of Skilled Labor - Narrow Sample

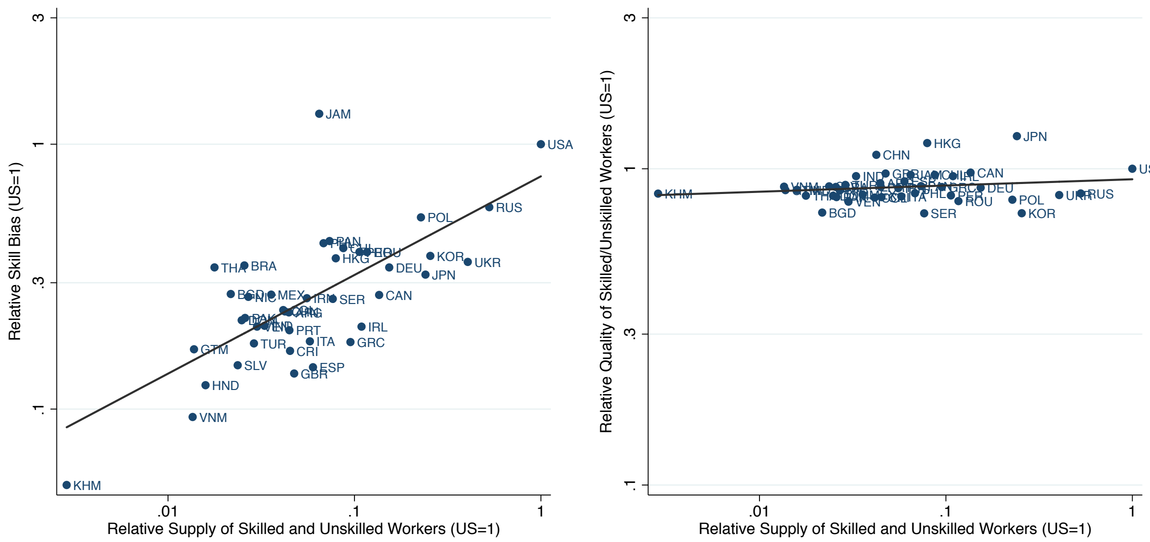
Notes: The figure plots on a log scale the relative efficiency and relative supply of skilled workers for countries in the narrow sample. Both variables are normalized so that they take value 1 for the United States. The solid line represents the best exponential fit.

Figure 3.4: Skill Premium and Skill Supply - Narrow Sample



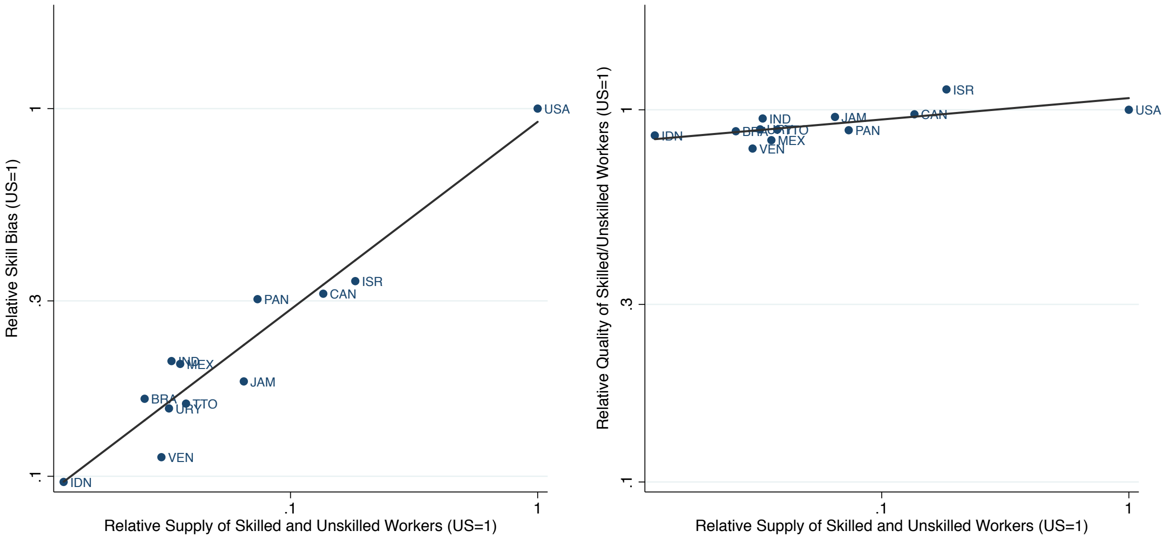
Notes: The figure plots the log skill premium and the log and relative supply of skilled workers for countries in the narrow sample. The solid line represents the best linear fit. The dashed line has the slope of the predicted relationship (-0.67) in a counterfactual where skill efficiency and supply are uncorrelated..

Figure 3.5: Technology Skill Bias, Skill Quality and Skill Supply - Broad Sample



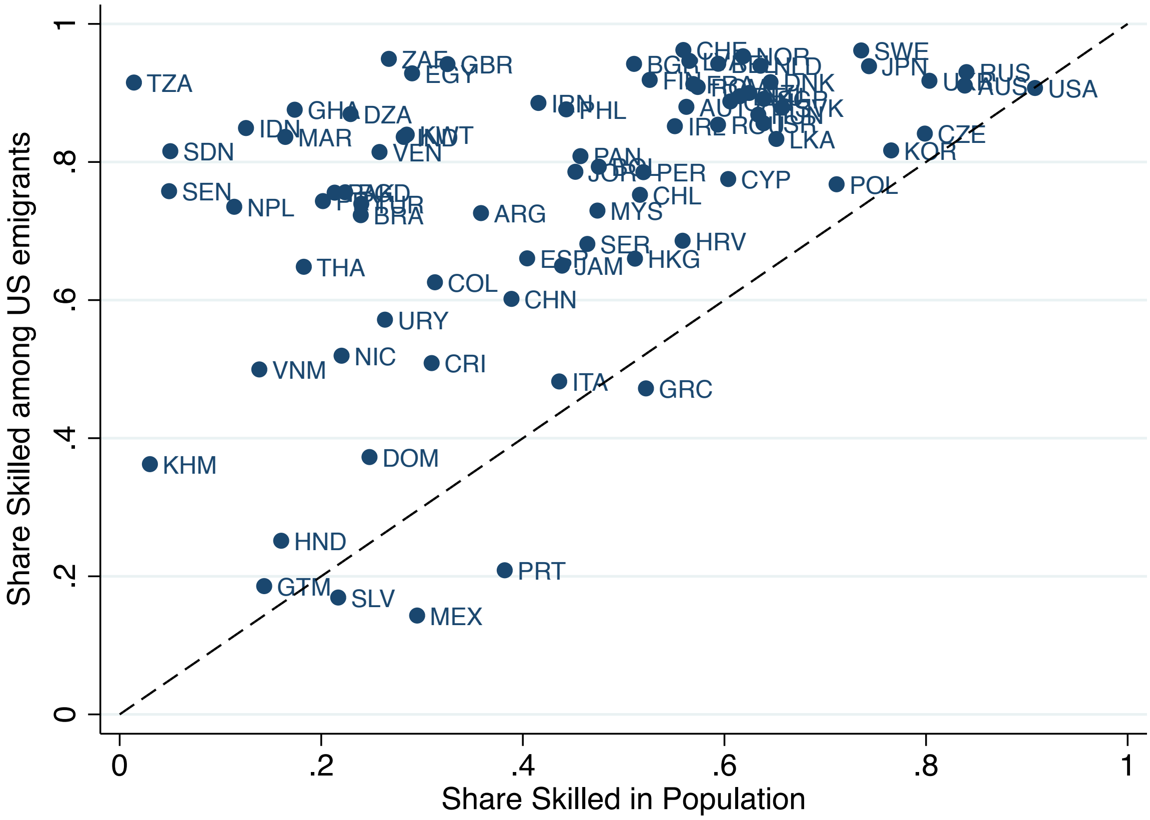
Notes: The left graph plots (on a log scale) the relative skill bias of technology again the relative supply of skilled labor. The right graph plots (on a log scale) the relative quality of skill labor again its relative supply. Only countries in the broad sample, with at least 100 unskilled and 100 skilled workers in the regression sample are included. All variables are normalised such that they take value 1 for the United States. The lines show the best exponential fits.

Figure 3.6: Technology Skill Bias, Skill Quality and Skill Supply - Narrow Sample



Notes: The left graph plots (on a log scale) the relative skill bias of technology again the relative supply of skilled labor. The right graph plots (on a log scale) the relative quality of skill labor again its relative supply. Only countries in the narrow sample are included. All variables are normalised such that they take value 1 for the United States. The lines show the best exponential fits.

Figure 3.7: Share of Skilled Workers among Emigrants by Country of Origin



Notes: The figure plots the share of skilled workers among emigrants to the US against the one in the country of origin. Only emigrants entirely educated in their country of origin are included. Skilled workers are defined as having completed secondary education or more.

Bibliography

Aaronson, Daniel, Lisa Barrow, and William Sander, "Teachers and Student Achievement in the Chicago Public High Schools," *Journal of Labor Economics*, 2007, *25*, 95–135.

Abramitzky, Ran, Leah Platt Boustan, and Katherine Eriksson, "Cultural Assimilation during the Age of Mass Migration," NBER Working Papers 22381, National Bureau of Economic Research July 2016.

Acemoglu, Daron, "Why Do New Technologies Complement Skills? Directed Technical Change And Wage Inequality," *The Quarterly Journal of Economics*, November 1998, *113* (4), 1055–1089.

_ , "Directed Technical Change," *Review of Economic Studies*, October 2002, *69* (4), 781–809.

_ **and Fabrizio Zilibotti**, "Productivity Differences," *The Quarterly Journal of Economics*, May 2001, *116* (2), 563–606.

Aguiar, Mark and Erik Hurst, "Measuring Trends in Leisure: The Allocation of Time over Five Decades," *Quarterly Journal of Economics*, 2007, *122* (3), 969–1006.

Albornoz, Facundo, Antonio Cabrales, and Esther Hauk, "Immigration and the school system," UC3M Working papers. Economics we1203, Universidad Carlos III de Madrid. Departamento de Economía January 2012.

Almond, Douglas and Janet Currie, *Human Capital Development before Age Five*, Vol. 4 of *Handbook of Labor Economics*, Elsevier,

Autor, David H. and David Dorn, "The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market," *American Economic Review*, 2013, *103* (5), 1553–97.

Barro, Robert J. and Jong Wha Lee, "A New Data Set of Educational Attainment in the World, 1950-2010," *Journal of Development Economics*, 2013, *104* (C), 184–198.

Bartik, Alex, "New International Data on Schools," Yale University Senior Thesis 2008.

Becker, Gary S., "Investment in Human Capital: A Theoretical Analysis," *Journal of Political Economy*, 1962, *70*, 9.

Belley, Philippe and Lance Lochner, "The Changing Role of Family Income and Ability in Determining Educational Achievement," *Journal of Human Capital*, 2007, *1* (1), 37–89.

Belot, Michèle V. K. and Timothy J. Hatton, "Immigrant Selection in the OECD," *The Scandinavian Journal of Economics*, 2012, *114* (4), 1105–1128.

Bils, Mark and Peter J. Klenow, "Does Schooling Cause Growth?," *American Economic Review*, December 2000, *90* (5), 1160–1183.

Bisin, Alberto and Thierry Verdier, "The Economics of Cultural Transmission and Socialization," NBER Working Papers 16512, National Bureau of Economic Research November 2010.

Black, Dan A. and Jeffrey A. Smith, "Estimating the Returns to College Quality with Multiple Proxies for Quality," *Journal of Labor Economics*, July 2006, *24* (3), 701–728.

Bleakley, Hoyt and Aimee Chin, "Age at Arrival, English Proficiency, and Social Assimilation among US Immigrants," *American Economic Journal: Applied Economics*, 2010, *2* (1), 165–92.

Borjas, George J., "Self-Selection and the Earnings of Immigrants," *The American Economic Review*, 1987, *77* (4), 531–553.

Borjas, George J, "Ethnic Capital and Intergenerational Mobility," *Quarterly Journal of Economics*, 1992, *107* (1), 123–50.

Bound, John and Sarah Turner, "Cohort crowding: How resources affect collegiate attainment," *Journal of Public Economics*, June 2007, *91* (5-6), 877–899.

Bowlus, Audra J. and Chris Robinson, "Human Capital Prices, Productivity, and Growth," *American Economic Review*, December 2012, *102* (7), 3483–3515.

Bratsberg, Bernt, "School Quality and Returns to Education of U.S. Immigrants," *Economic Inquiry*, April 2002, *40* (2), 177–198.

Brown, Meta, John Karl Scholz, and Ananth Seshadri, "A New Test of Borrowing Constraints for Education," *Review of Economic Studies*, 2012, *79* (2), 511–538.

Brueckner, Markus and Daniel Lederman, "Effects of income inequality on aggregate output," Policy Research Working Paper Series 7317, The World Bank June 2015.

Buera, Francisco J., Joseph P. Kaboski, and Yongseok Shin, "Finance and Development: A Tale of Two Sectors," *American Economic Review*, August 2011, *101* (5), 1964–2002.

Cabrera, A. F. and S. La Nasa, "On the path to college: Three critical tasks facing America's disadvantaged.," *Research in Higher Education*, 2001, *42* (2).

Cameron, Stephen V. and James J. Heckman, "Life Cycle Schooling and Dynamic Selection Bias: Models and Evidence for Five Cohorts of American Males," *Journal of Political Economy*, April 1998, *106*, 262–333.

Card, David, "Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems," *Econometrica*, September 2001, *69* (5), 1127–1160.

_ , Jörg Heining, and Patrick Kline, "Workplace Heterogeneity and the Rise of West German Wage Inequality," *Quarterly Journal of Economics*, 2013, *128* (3), 967–1015.

Carneiro, Pedro and James J. Heckman, "The Evidence on Credit Constraints in Post–secondary Schooling," *Economic Journal*, October 2002, *112* (482), 705–734.

Carrell, Scott E. and Bruce Sacerdote, "Late Interventions Matter Too: The Case of College Coaching New Hampshire," NBER Working Papers 19031, National Bureau of Economic Research, Inc May 2013.

Cascio, Elizabeth U., "School Progression and the Grade Distribution of Students: Evidence from the Current Population Survey," IZA Discussion Papers 1747, Institute for the Study of Labor September 2005.

— **and Ethan G. Lewis**, "Schooling and the AFQT: Evidence from School Entry Laws," NBER Working Papers 11113, National Bureau of Economic Research, Inc February 2005.

Caselli, Francesco, *Technology Differences over Space and Time*, Princeton University Press, 2016.

— **and Nicola Gennaioli**, "Dynastic Management," *Economic Inquiry*, 01 2013, *51* (1), 971–996.

— **and Wilbur John Coleman**, "The World Technology Frontier," *American Economic Review*, June 2006, *96* (3), 499–522.

— **, Jacopo Ponticelli, and Federico Rossi**, *A New Data Set on Mincerian Returns* Technology Differences across Space and Time, Princeton University Press, 2016.

Chiquiar, Daniel and Gordon H. Hanson, "International Migration, Self-Selection, and the Distribution of Wages: Evidence from Mexico and the United States," *Journal of Political Economy*, 2005, *113* (2), 239–281.

Chiswick, Barry R, "The Effect of Americanization on the Earnings of Foreign-born Men," *Journal of Political Economy*, October 1978, *86* (5), 897–921.

Chiswick, Barry R., "Are Immigrants Favorably Self-Selected? An Economic Analysis," IZA Discussion Papers 131, Institute for the Study of Labor March 2000.

Chua, Amy, *Battle Hymn of the Tiger Mother*, The Penguin Press, 2011.

Ciccone, Antonio and Giovanni Peri, "Long-Run Substitutability Between More and Less Educated Workers: Evidence from U.S. States, 1950-1990," *The Review of Economics and Statistics*, November 2005, *87* (4), 652–663.

— **and** — , "Identifying Human-Capital Externalities: Theory with Applications," *Review of Economic Studies*, 2006, *73* (2), 381–412.

Coleman, Margo and Thomas DeLeire, "An Economic Model of Locus of Control and the Human Capital Investment Decision," *Journal of Human Resources*, 2003, *38* (3).

Cunha, Flavio, James J. Heckman, and Susanne M. Schennach, "Estimating the Technology of Cognitive and Noncognitive Skill Formation," *Econometrica*, 05 2010, *78* (3), 883–931.

Doepke, Matthias and Fabrizio Zilibotti, "Parenting with style: Altruism and paternalism in intergenerational preference transmission," *Econometrica*, 2017.

Dohmen, Thomas, Benjamin Enke, Armin Falk, David Huffman, and Uwe Sunde, "Patience and the Wealth of Nations," Working Papers 2016-012, Human Capital and Economic Opportunity Working Group April 2016.

Dronkers, Jaap and Manon de Heus, *Educational Performance of the Children of Immigrants in Sixteen OECD Countries* A World in Motion: Trends in Migration and Migration Policy, Oxford University Press,

Duckworth, Angela L. and Martin E. P. Seligman, "Self-discipline outdoes IQ in predicting academic performance of adolescents," *Psychological Science*, 2005, *16* (12), 939–944.

Dustmann, Christian, Tommaso Frattini, and Gianandrea Lanzara, "Educational Achievement of Second-generation Immigrants: an International Comparison," *Economic Policy*, 2012, *27* (69), 143–185.

Ellwood, David and Thomas J. Kane, "Who is Getting a College Education: Family Background and the Growing Gaps in Enrollment," in Sheldon Danziger and Jane Waldfogel, eds., *Securing the Future*, Russell Sage Foundation, 2000.

Feliciano, Cynthia, "Does Selective Migration Matter? Explaining Ethnic Disparities in Educational Attainment among Immigrants' Children," *International Migration Review*, 2005, *39* (4), 841–871.

_, "Educational selectivity in U.S. Immigration: How do immigrants compare to those left behind?," *Demography*, 2005, *42* (1), 131–152.

Fernandez, Raquel, "Does Culture Matter?," in "Handbook of Social Economics," Elsevier, 2011.

_ **and Alessandra Fogli**, "Culture: An Empirical Investigation of Beliefs, Work, and Fertility," *American Economic Journal: Macroeconomics*, 2009, *1* (1), 146–77.

Figlio, David, Paola Giuliano, Umut Özek, and Paola Sapienza, "Long-Term Orientation and Educational Performance," NBER Working Papers 22541, National Bureau of Economic Research August 2016.

Finkelstein, Amy, Matthew Gentzkow, and Heidi Williams, "Sources of Geographic Variation in Health Care: Evidence From Patient Migration," *The Quarterly Journal of Economics*, 2016, *131* (4), 1681–1726.

Frazis, Harley, "Selection Bias and the Degree Effect," *Journal of Human Resources*, 1993, *28* (3), 538–554.

Furtado, Delia and Stephen J. Trejo, "Interethnic marriages and their economic effects," in "International Handbook on the Economics of Migration" Chapters, Edward Elgar Publishing, 2013, chapter 15, pp. 276–292.

Galor, Oded and Ömer Özak, "The Agricultural Origins of Time Preference," *American Economic Review*, October 2016, *106* (10), 3064–3103.

Ganzeboom, H.B.G., P.M. de Graaf, D.J. Treiman, and J. de Leeuw, "A standard international socio-economic index of occupational status," WORC Paper, Tilburg University, Work and Organization Research Centre 1992.

Giuliano, Paola, "Living Arrangements in Western Europe: Does Cultural Origin Matter?," *Journal of the European Economic Association*, 2007, *5* (5), 927–952.

Glewwe, Paul and Michael Kremer, *Schools, Teachers, and Education Outcomes in Developing Countries*, Vol. 2 of *Handbook of the Economics of Education*, Elsevier, June

Gorard, Stephen, Beng Huat See, and Peter Davies, "The impact of attitudes and aspirations on educational attainment and participation," Technical Report, Joseph Rowntree Foundation 2012.

Gordon, Milton, *Assimilation in American Life*, Oxford University Press, 1964.

Grogger, Jeffrey and Gordon H. Hanson, "Income maximization and the selection and sorting of international migrants," *Journal of Development Economics*, May 2011, *95* (1), 42–57.

Hall, Robert E. and Charles I. Jones, "Why Do Some Countries Produce So Much More Output per Worker than Others?," *The Quarterly Journal of Economics*, February 1999, *114* (1), 83–116.

Hanushek, Eric A. and Ludger Woessmann, "Do Better Schools lead to More Growth? Cognitive Skills, Economic Outcomes, and Causation," *Journal of Economic Growth*, 2012, *17* (4), 267–321.

_ and _ , "Schooling, educational achievement, and the Latin American growth puzzle," *Journal of Development Economics*, 2012, *99* (2), 497–512.

_ , Charles Ka Yui Leung, and Kuzey Yilmaz, "Borrowing Constraints, College Aid, and Intergenerational Mobility," *Journal of Human Capital*, 2014, *8* (1), 1 – 41.

Heckman, James and Flavio Cunha, "The Technology of Skill Formation," *American Economic Review*, May 2007, *97* (2), 31–47.

Heckman, James J., Jora Stixrud, and Sergio Urzua, "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior," *Journal of Labor Economics*, July 2006, *24* (3), 411–482.

Hendricks, Lutz and Todd Schoellman, "Student abilities during the expansion of US education," *Journal of Monetary Economics*, 2014, *63* (C), 19–36.

_ and _ , "Human capital and development accounting: New evidence from wage gains at migration," CFS Working Paper Series 557, Center for Financial Studies 2016.

Hill, C. B. and G. C. Winston, "Low-income students and highly selective private colleges: Geography, searching, and recruiting," *Economics of Education Review*, 2010, *29* (4).

Hofstede, Geert, *Cultures and organizations: software of the mind*, McGraw-Hill, 1991.

_ , Gert Jan Hofstede, and Michael Minkov, *Cultures and organizations: software of the mind: intercultural cooperation and its importance for survival*, McGraw-Hill, 2010.

Hoxby, Caroline M. and Christopher Avery, "The Missing One-Offs: The Hidden Supply of High Achieving, Low Income Students," NBER Working Papers 18586, National Bureau of Economic Research December 2012.

_ and Sarah Turner, "Expanding College Opportunities for High-Achieving, Low Income Students," SIEPR Discussion Paper 12-014, Stanford Institute for Economic Policy Research 2013.

Hsieh, Chang-Tai and Peter J. Klenow, "Misallocation and Manufacturing TFP in China and India," *The Quarterly Journal of Economics*, November 2009, *124* (4), 1403–1448.

_ , Erik Hurst, Charles I. Jones, and Peter J. Klenow, "The Allocation of Talent and U.S. Economic Growth," NBER Working Papers 18693, National Bureau of Economic Research Jan 2013.

Isphording, Ingo E., Marc Piopiunik, and Núria Rodríguez-Planas, "Speaking in numbers: The effect of reading performance on math performance among immigrants," *Economics Letters*, 2016, *139* (C), 52–56.

Jaeger, David A and Marianne E Page, "Degrees Matter: New Evidence on Sheepskin Effects in the Returns to Education," *The Review of Economics and Statistics*, November 1996, *78* (4), 733–40.

Jerrim, John, "Why do East Asian children perform so well in PISA? An investigation of Western-born children of East Asian descent," *Oxford Review of Education*, 2015, *41* (3), 310–333.

Jones, Benjamin F., "The Human Capital Stock: A Generalized Approach," *American Economic Review*, November 2014, *104* (11), 3752–3777.

Keane, Michael P and Kenneth I Wolpin, "The Effect of Parental Transfers and Borrowing Constraints on Educational Attainment," *International Economic Review*, November 2001, *42* (4), 1051–1103.

Kinsler, Josh and Ronni Pavan, "Family Income and Higher Education Choices: The Importance of Accounting for College Quality," *Journal of Human Capital*, 2011, *5* (4), 453 – 477.

Klenow, Peter and Andrés Rodríguez-Clare, "The Neoclassical Revival in Growth Economics: Has It Gone Too Far?," in "NBER Macroeconomics Annual 1997, Volume 12" NBER Chapters, National Bureau of Economic Research, May 1997, pp. 73–114.

Lagakos, David, Benjamin Moll, Tommaso Porzio, Nancy Qian, and Todd Schoellman, "Life-Cycle Human Capital Accumulation Across Countries: Lessons From U.S. Immigrants," NBER Working Papers 21914, National Bureau of Economic Research January 2016.

Levels, Mark, Jaap Dronkers, and Gerbert Kraaykamp, "Immigrant Children's Educational Achievement in Western Countries: Origin, Destination, and Community Effects on Mathematical Performance," *American Sociological Review*, 2008, *73* (5), 835–853.

Lichter, Daniel T. and Zhenchao Qian, "Measuring Marital Assimilation: Intermarriage Among Natives and Immigrants," *Social Science Research*, 2001, *30*, 289–312.

Little, Roderick and Donald Rubin, *Statistical Analysis with Missing Data*, Wiley & Sons, 1987.

Malmberg, Hannes, "Human Capital and Development Accounting Revisited," Working Paper 2017.

Manuelli, Rodolfo E. and Ananth Seshadri, "Human Capital and the Wealth of Nations," *American Economic Review*, 2014, *104* (9), 2736–62.

Mayer, Thierry and Soledad Zignago, "Notes on CEPII's distances measures: The GeoDist database," Working Papers 2011-25, CEPII Research Center December 2011.

McKenzie, David and Hillel Rapoport, "Self-Selection Patterns in Mexico-U.S. Migration: The Role of Migration Networks," *The Review of Economics and Statistics*, November 2010, *92* (4), 811–821.

Meng, Xin and Robert G. Gregory, "Intermarriage and the Economic Assimilation of Immigrants," *Journal of Labor Economics*, 2005, *23* (1), 135–176.

Moriconi, Simone and Giovanni Peri, "Country-Specific Preferences and Employment Rates in Europe," NBER Working Papers 21561, National Bureau of Economic Research September 2015.

Nielsen, Helena Skyt and Beatrice Schindler Rangvid, "The Impact of Parents' Years since Migration on Children's Academic Achievement," IZA Discussion Papers 6242, Institute for the Study of Labor December 2011.

NLS, *NLSY79 Profiles of American Youth: Addendum to Attachment 106* Columbus: Ohio State University 1992.

OECD, "PISA Data Analysis Manual: SPSS," Technical Report 2009.

_ , "PISA 2012 Results: What Students Know and Can Do," Technical Report 2012.

_ , "PISA 2012 Technical Report," Technical Report 2012.

Oreopoulos, Philip and Marianne E. Page, "The Intergenerational Effects of Compulsory Schooling," *Journal of Labor Economics*, 2006, *24* (4), 729–760.

Ozden, Cağlar, Christopher R. Parsons, Maurice Schiff, and Terrie L. Walmsley, "Where on Earth is Everybody? The Evolution of Global Bilateral Migration 1960–2000," *World Bank Economic Review*, 2011, *25* (1), 12–56.

Pagnini, Deanna and Philip Morgan, "Intermarriage and Social Distance Among U.S. Immigrants at the Turn of the Century," *American Journal of Sociology*, 1990, *96* (2), 405–432.

Pallais, A. and S. Turner, "Opportunities for low-income students at top colleges and universities: Policy initiatives and the distribution of students," *National Tax Journal*, 2006, *59* (2).

Park, Jin Heum, "Estimation of sheepskin effects using the old and the new measures of educational attainment in the Current Population Survey," *Economics Letters*, February 1999, *62* (2), 237–240.

Psacharopoulos, George and Harry Anthony Patrinos, "Returns to investment in education: a further update," *Education Economics*, 2004, *12* (2), 111–134.

Ramey, Valerie, "Is there a "Tiger Mother" Effect? Time Use Across Ethnic Groups," *Economics in Action*, 2011, *4.*

Roy, A. D., "Some Thoughts on the Distribution of Earnings," *Oxford Economic Papers*, 1951, *3* (2), 135–146.

Rubinstein, Yona and James J. Heckman, "The Importance of Noncognitive Skills: Lessons from the GED Testing Program," *American Economic Review*, May 2001, *91* (2), 145–149.

Schoellman, Todd, "Education Quality and Development Accounting," *Review of Economic Studies*, 2012, *79* (1), 388–417.

_ , "Early Childhood Human Capital and Development," *American Economic Journal: Macroeconomics*, July 2016, *8* (3), 145–174.

Smith, Jonathan, Matea Pender, and Jessica Howell, “The full extent of student-college academic undermatch,” *Economics of Education Review*, 2013, *32* (C), 247–261.

Spolaore, Enrico and Romain Wacziarg, “Ancestry, Language and Culture,” CESifo Working Paper Series 5388, CESifo Group Munich 2015.

Søren, Eric W. Holman Wichmann and Cecil H. Brown, “The ASJP Database (version 17),” Technical Report 2016.

Vollrath, Dietrich, “The efficiency of human capital allocations in developing countries,” *Journal of Development Economics*, 2014, *108* (C), 106–118.

Woessmann, Ludger, “The Importance of School Systems: Evidence from International Differences in Student Achievement,” *Journal of Economic Perspectives*, 2016, *30* (3), 3–32.

Wolfe, Raymond N. and Scott D. Johnson, “Personality as a predictor of college performance,” *Educational and Psychological Measurement*, 1995, *55* (2), 177–185.

Appendices

Appendix A

Appendices to Chapter 1

A.1 Data Appendix

A.1.1 Data Construction

Given that individual host countries have great flexibility in choosing how to report parents' countries of birth, some aggregation is necessary to get a set of countries consistently defined over time. For what concerns countries included in the PISA sample, we make the following adjustments: we code *Yugoslavia* and similar labels as Serbia and Montenegro, *USSR* and similar labels as Russia, *Albania or Kosovo* as Albania, *France or Belgium* as France, *Germany or Austria* as Germany, *China (including Hong Kong)* as China. Moreover, for the purpose of estimating (1.4), we group countries of origin not belonging to the PISA sample in several categories (introducing a fixed effect for each of those): in particular, we create dummies for individual countries when possible (Belarus, Bolivia, Bosnia, Pakistan, Paraguay, Philippines, Ukraine), aggregate others in broad geographical groups (Africa, Europe, Middle East) and classify any remaining case as Rest of the World. We drop all observations with inconsistent or missing information on students' or parents' countries of birth.

Parents' educational attainment is reported according to the ISCED 1997 classification system. We group levels 0 and 1 into *primary* education, levels 2, 3 and 4 into *secondary* education and levels 5 and 6 into *tertiary* education.

A.1.2 Additional Summary Statistics

Table A.1: Average PISA Scores across Regions

	Math	Reading	Science	# Countries
China	1.33	0.96	1.07	1
Other East Asia	0.79	0.58	0.67	6
Canada	0.63	0.66	0.68	1
EU North	0.57	0.53	0.58	15
Oceania	0.54	0.62	0.67	2
US	0.26	0.45	0.43	1
EU South	0.13	0.18	0.21	5
EU East	-0.08	-0.16	-0.06	19
Other Asia	-0.42	-0.38	-0.36	5
Middle East/NA	-0.55	-0.40	-0.43	7
Latin America	-0.58	-0.38	-0.46	11

Notes: The Table shows the average PISA score of native students across countries belonging to each region, for all available waves (for Science, only waves from 2006 onwards are considered, since the scale was established in 2006 and results from 2003 are not fully comparable with the subsequent ones). Country averages are computed using the provided sample weights. Scores are standardized to have mean 0 and (individual-level) standard deviation 1 across the (pooled, equally weighted) 73 countries participating to at least one wave of the test. Countries are assigned to regional groups as follows. *East Asia*: China, Hong Kong, Japan, Macao, Singapore, South Korea, Taiwan. *EU North*: Austria, Belgium, Denmark, Finland, France, Germany, Iceland, Ireland, Liechtenstein, Luxembourg, Netherlands, Norway, Sweden, Switzerland, United Kingdom. *Oceania*: Australia, New Zealand. *EU South*: Greece, Italy, Malta, Portugal, Spain. *EU East*: Albania, Azerbaijan, Bulgaria, Croatia, Czechia, Estonia, Georgia, Hungary, Kazakhstan, Kyrgyzstan, Latvia, Lithuania, Moldova, Poland, Romania, Russia, Serbia and Montenegro, Slovakia, Slovenia. *Other Asia*: India, Indonesia, Malaysia, Thailand, Vietnam. *Latin America*: Argentina, Brazil, Chile, Colombia, Costa Rica, Mexico, Panama, Peru, Trinidad and Tobago, Uruguay, Venezuela. *Middle East / North Africa*: Israel, Jordan, Mauritius, Qatar, Tunisia, Turkey, United Arab Emirates.

Table A.2: Second Generation Immigrants by Country of Origin - PISA

Country of Origin	Mothers Number	Mothers # Host Countries	Mothers Top Host Country	Fathers Number	Fathers # Host Countries	Fathers Top Host Country
Albania	330	5	Greece (183)	306	5	Greece (162)
Argentina	69	2	Uruguay (68)	64	1	Uruguay (64)
Australia	150	2	New Zealand (149)	117	1	New Zealand (117)
Austria	239	2	Switzerland (181)	175	2	Switzerland (137)
Belgium	258	3	Luxembourg (244)	239	2	Luxembourg (219)
Brazil	187	4	Uruguay (78)	162	4	Uruguay (72)
Bulgaria	30	1	Turkey (30)	16	1	Turkey (16)
Canada	2	1	Ireland (2)	2	1	Ireland (2)
Chile	58	1	Argentina (58)	48	1	Argentina (48)
China	14161	12	Macao (9149)	13529	11	Macao (8416)
Colombia	5	1	Costa Rica (5)	5	1	Costa Rica (5)
Croatia	190	3	Serbia-Mont. (101)	167	3	Serbia-Mont. (81)
Czech Republic	187	2	Slovakia (179)	176	2	Slovakia (167)
Denmark	75	2	Norway (74)	91	1	Norway (91)
Estonia	81	1	Finland (81)	53	1	Finland (53)
France	1312	7	Switzerland (626)	1110	7	Switzerland (456)
Germany	1364	9	Switzerland (630)	1086	9	Switzerland (461)
Greece	85	2	Australia (68)	133	2	Australia (112)
Hong Kong	236	2	Macao (168)	432	3	Macao (348)
Hungary	16	2	Austria (14)	15	2	Austria (12)
India	218	4	Australia (189)	220	4	Australia (187)
Italy	1510	9	Switzerland (1006)	2606	9	Switzerland (1741)
Jordan	155	1	Qatar (155)	119	1	Qatar (119)
Liechtenstein	38	1	Switzerland (38)	27	1	Switzerland (27)
Macao	138	1	Hong Kong (138)	123	1	Hong Kong (123)
Malaysia	65	4	Australia (53)	55	4	Australia (45)
Netherlands	226	5	Belgium (195)	264	4	Belgium (192)
New Zealand	781	1	Australia (781)	790	1	Australia (790)
Panama	9	1	Costa Rica (9)	15	1	Costa Rica (15)
Poland	275	3	Germany (211)	213	3	Germany (173)
Portugal	2646	4	Luxembourg (1762)	2517	5	Luxembourg (1730)
Romania	49	2	Austria (47)	55	3	Austria (43)
Russia	4216	13	Estonia (1219)	4092	13	Estonia (1225)
Serbia-Mont.	2615	9	Switzerland (1548)	2653	9	Switzerland (1563)
Singapore	8	1	Indonesia (8)	8	2	Indonesia (7)
Slovakia	472	2	Czech Republic (467)	569	2	Czech Republic (564)
Slovenia	11	2	Austria (8)	17	2	Austria (10)
South Korea	42	2	Australia (30)	43	2	Australia (33)
Spain	334	5	Switzerland (317)	409	4	Switzerland (391)
Sweden	362	2	Finland (230)	272	2	Finland (173)
Switzerland	106	1	Liechtenstein (106)	90	1	Liechtenstein (90)
Taiwan	26	1	Hong Kong (26)	9	2	Hong Kong (6)
Thailand	13	1	Finland (13)	2	1	Finland (2)
Turkey	2411	8	Denmark (535)	2648	8	Switzerland (591)
United Kingdom	3514	5	Australia (2142)	3659	5	Australia (2313)
United States	407	5	Mexico (198)	532	5	Mexico (326)
Uruguay	79	1	Argentina (79)	72	1	Argentina (72)
Vietnam	304	4	Australia (249)	299	3	Australia (240)
Average	834.69	3.38		839.67	3.31	

Notes: The Table shows summary statistics on second generation immigrants from each country of origin in the PISA sample (with at least one observation per parent). *# Host Countries* is the number of different host countries in which second generation immigrants are observed. *Top Host Country* is the host country where the highest number (reported in brackets) of second generation immigrants are observed.

Table A.3: Second Generation Immigrants by Host Country - PISA

	Mothers			Fathers		
Host Country	Number	# Countries of Origin	Top Country of Origin (in PISA)	Number	# Countries of Origin	Top Country of Origin (in PISA)
Argentina	541	6	Uruguay (79)	497	6	Uruguay (72)
Australia	8403	17	United Kingdom (2142)	8740	17	United Kingdom (2313)
Austria	1623	15	Turkey (370)	1603	15	Turkey (393)
Belgium	2820	7	Turkey (384)	3178	7	Turkey (437)
Costa Rica	423	3	Panama (9)	490	3	Panama (15)
Croatia	2013	4	Serbia-Mont. (347)	1819	4	Serbia-Mont. (329)
Czech Republic	661	6	Slovakia (467)	872	6	Slovakia (564)
Denmark	2343	6	Turkey (535)	2460	6	Turkey (541)
Estonia	1492	2	Russia (1219)	1605	2	Russia (1225)
Finland	1045	10	Sweden (230)	1193	10	Sweden (173)
Georgia	58	2	Russia (40)	47	2	Russia (32)
Germany	1194	10	Turkey (392)	1247	10	Turkey (424)
Greece	1167	3	Russia (198)	706	3	Albania (162)
Hong Kong	5110	4	China (4470)	4982	4	China (4650)
Indonesia	39	5	Singapore (8)	41	3	Singapore (7)
Ireland	1085	16	United Kingdom (872)	971	15	United Kingdom (761)
Israel	2015	5	Russia (531)	2130	5	Russia (520)
Kazakhstan	1106	2	Russia (926)	1056	2	Russia (864)
Kyrgyzstan	443	2	Russia (98)	275	2	Russia (94)
Latvia	1932	4	Russia (811)	2197	4	Russia (919)
Liechtenstein	301	11	Switzerland (106)	258	11	Switzerland (90)
Luxembourg	4116	10	Portugal (1762)	4207	10	Portugal (1730)
Macao	9755	5	China (9149)	9238	7	China (8416)
Mauritius	80	4	China (11)	54	4	China (8)
Mexico	972	4	United States (198)	1256	4	United States (326)
Moldova	174	3	Russia (53)	159	4	Russia (50)
Netherlands	1468	16	Turkey (167)	1547	16	Turkey (191)
New Zealand	1736	8	United Kingdom (471)	1869	8	United Kingdom (548)
Norway	1025	3	Sweden (132)	1013	3	Sweden (99)
Portugal	1450	5	Brazil (58)	1232	5	Brazil (57)
Qatar	4587	4	Jordan (155)	4027	4	Jordan (119)
Serbia-Mont.	2088	4	Croatia (101)	1617	4	Croatia (81)
Slovakia	524	3	Czech Republic (179)	526	3	Czech Republic (167)
Slovenia	1591	3	Italy (7)	1630	3	Italy (9)
South Korea	26	5	China (10)	-	-	-
Switzerland	8036	11	Serbia-Mont. (1548)	7886	11	Italy (1741)
Turkey	211	5	Germany (62)	176	5	Germany (29)
United Kingdom	2025	7	China (23)	2181	7	China (24)
Uruguay	265	4	Brazil (78)	283	4	Brazil (72)
Average	1947.26	6.26		1930.03	6.15	

Notes: The Table shows summary statistics on second generation immigrants observed in each country in the PISA sample, across all available waves. Only host countries with second generation immigrants from at least one country of origin in the PISA sample are included. *# Countries of Origin* is the number of different countries of origin of second generation immigrants in a given host country. *Top Country of Origin (in PISA)* is the country of origin from which the highest number (across all countries in the PISA sample, not considering other countries of origin) of second generation immigrants in a given host country are observed (number reported in brackets).

A.2 Robustness of Baseline Result

A.2.1 PISA

A.2.1.1 Results for Second Generation Immigrants on the Father's Side

Table A.4: Main Results for Fathers - PISA

	Dependent Variable: Math Test Score					
	[1]	[2]	[3] All	[4]	[5]	[6] No East Asia
Score Country f	0.676***	0.535***	0.230***	0.225***	0.212***	0.179*
	(0.079)	(0.166)	(0.086)	(0.080)	(0.072)	(0.108)
Female	-0.139***	-0.140***	-0.204***	-0.200***	-0.197***	-0.181***
	(0.028)	(0.024)	(0.028)	(0.028)	(0.028)	(0.035)
Father Sec Edu				-0.003	-0.016	-0.025
				(0.023)	(0.022)	(0.041)
Father Ter Edu				0.080**	0.030	0.016
				(0.034)	(0.033)	(0.046)
Mother Sec Edu				0.025	0.005	0.056
				(0.044)	(0.043)	(0.082)
Mother Ter Edu				0.073	0.025	0.067
				(0.056)	(0.057)	(0.103)
Mother Working $\times$ ISEI					0.003***	0.003***
					(0.001)	(0.001)
Father Working $\times$ ISEI					0.003***	0.003***
					(0.001)	(0.001)
N	40304	40304	40304	40304	40304	26160
# Country f	48	48	48	48	48	42
R Squared	0.16	0.25	0.66	0.66	0.66	0.63
Host Country FE	No	Yes	Yes	Yes	Yes	Yes
School FE	No	No	Yes	Yes	Yes	Yes

Notes: The Table shows results for second generation immigrants on the father's side. Sample includes only cases where both parents report a country of origin and the country of origin of the father runs a PISA test on natives. *Score Country f* is the average math PISA score of natives (standardized to have mean 0 and standard deviation 1 across all countries participating to the test) in the country of birth of the father, across all available waves. All specifications control for intercept, students' age (in months), wave fixed effect and a dummy for mother's immigrant status; specifications 5-6 additionally control for dummies for parents' employment status (full-time employed, part-time employed, not working). *Working* refers to either full-time or part-time employed. Observations are weighted according to the provided sample weights. Standard errors are clustered by father's country of origin, and inflated by the estimated measurement error in test scores. * denotes significance at 10%, ** at 5%, *** at 1%.

A.2.1.2 Results for Second Generation Immigrants and Natives

Table A.5: Main Results for Mothers and Fathers - PISA

	Dependent Variable: Math Test Score					
	[1]	[2]	[3] All	[4]	[5]	[6] No East Asia
Score Country *m*	0.307***	0.304***	0.163***	0.156***	0.149***	0.132**
	(0.106)	(0.118)	(0.042)	(0.041)	(0.042)	(0.058)
Score Country *f*	0.400***	0.386***	0.202***	0.195***	0.184***	0.172***
	(0.098)	(0.095)	(0.041)	(0.039)	(0.038)	(0.067)
Score Country *m* * Native Mother	0.219**	0.103	0.026	0.024	0.014	0.050
	(0.087)	(0.071)	(0.030)	(0.033)	(0.034)	(0.049)
Score Country *f* * Native Father	0.093	-0.032	-0.045	-0.053	-0.060*	-0.042
	(0.099)	(0.091)	(0.036)	(0.037)	(0.034)	(0.060)
Female	-0.117***	-0.117***	-0.151***	-0.146***	-0.143***	-0.143***
	(0.011)	(0.011)	(0.012)	(0.013)	(0.012)	(0.013)
Native Mother	-0.038	0.003	0.015	0.006	-0.002	0.031
	(0.060)	(0.054)	(0.024)	(0.036)	(0.038)	(0.031)
Native Father	0.016	0.071	0.057**	0.034	0.026	0.025
	(0.067)	(0.060)	(0.023)	(0.025)	(0.025)	(0.030)
Father Sec Edu				0.015	0.000	0.009
				(0.022)	(0.022)	(0.035)
Father Ter Edu				0.080**	0.024	0.024
				(0.036)	(0.032)	(0.041)
Mother Sec Edu				0.025	0.012	0.056*
				(0.032)	(0.035)	(0.030)
Mother Ter Edu				0.066*	0.014	0.050
				(0.039)	(0.039)	(0.042)
Native Father × Father Sec Edu				0.023	0.023	0.014
				(0.024)	(0.023)	(0.036)
Native Father × Father Ter Edu				0.024	0.020	0.018
				(0.041)	(0.037)	(0.047)
Native Mother × Mother Sec Edu				0.002	0.000	-0.047
				(0.036)	(0.037)	(0.030)
Native Mother × Mother Ter Edu				0.019	0.015	-0.024
				(0.045)	(0.045)	(0.045)
Mother Working × ISEI					0.004***	0.004***
					(0.000)	(0.000)
Father Working × ISEI					0.004***	0.004***
					(0.001)	(0.001)
N	1181347	1181347	1181347	1181347	1181347	1089297
# Country *m*	49	49	49	49	49	42
# Country *f*	48	48	48	48	48	42
R Squared	0.34	0.34	0.61	0.61	0.62	0.60
Host Country FE	No	Yes	Yes	Yes	Yes	Yes
School FE	No	No	Yes	Yes	Yes	Yes

Notes: The Table shows results for second generation immigrants and natives. The sample includes only cases where both parents report a country of origin that runs a PISA test on natives. *Score Country m* and *Score Country f* are the average math PISA score of natives (standardized to have mean 0 and standard deviation 1 across all countries participating to the test) in the country of birth of the mother and father, across all available waves. All specifications control for intercept, students' age (in months), wave fixed effect and a dummy for father's immigrant status; specifications 5-6 additionally control for dummies for parents' employment status (full-time employed, part-time employed, not working). *Working* refers to either full-time or part-time employed. Observations are weighted according to the provided sample weights. Standard errors are clustered by mother's and father's country of origin, and inflated by the estimated measurement error in test scores. * denotes significance at 10%, ** at 5%, *** at 1%.

A.2.1.3 Results for Reading and Science

Table A.6: Results for PISA Reading

	Dependent Variable: Reading Test Score					
	[1]	[2]	[3]	[4]	[5]	[6]
			All			No East Asia
Score Read Country *m*	0.724***	0.333*	0.165***	0.166***	0.157***	0.119**
	(0.146)	(0.184)	(0.063)	(0.058)	(0.056)	(0.058)
Female	0.311***	0.301***	0.240***	0.243***	0.246***	0.269***
	(0.038)	(0.035)	(0.023)	(0.024)	(0.024)	(0.017)
Father Sec Edu				0.068	0.055	0.110**
				(0.048)	(0.049)	(0.055)
Father Ter Edu				0.145***	0.100*	0.145**
				(0.056)	(0.059)	(0.068)
Mother Sec Edu				-0.002	-0.017	-0.002
				(0.028)	(0.029)	(0.055)
Mother Ter Edu				0.044	0.002	0.010
				(0.045)	(0.046)	(0.075)
Mother Working $\times$ ISEI					0.002***	0.003***
					(0.001)	(0.001)
Father Working $\times$ ISEI					0.003***	0.003***
					(0.000)	(0.001)
N	40067	40067	40067	40067	40067	25454
# Country *m*	49	49	49	49	49	42
R Squared	0.14	0.26	0.69	0.69	0.70	0.67
Host Country FE	No	Yes	Yes	Yes	Yes	Yes
School FE	No	No	Yes	Yes	Yes	Yes

Notes: The Table shows results for second generation immigrants on the mother's side. The sample includes only cases where both parents report a country of origin and the country of origin of the mother runs a PISA test on natives. *Score Read Country m* is the average reading PISA score of natives (standardized to have mean 0 and standard deviation 1 across all countries participating to the test) in the country of birth of the mother, across all available waves. All specifications control for intercept, students' age (in months), wave fixed effect and a dummy for father's immigrant status. Observations are weighted according to the provided sample weights. Standard errors are clustered by mother's country of origin, and inflated by the estimated measurement error in test scores. * denotes significance at 10%, ** at 5%, *** at 1%.

Table A.7: Results for PISA Science

	Dependent Variable: Science Test Score					
	[1]	[2]	[3]	[4]	[5]	[6]
			All			No East Asia
Score Science Country *m*	0.783***	0.537***	0.280***	0.273***	0.262***	0.253***
	(0.113)	(0.196)	(0.087)	(0.081)	(0.075)	(0.096)
Female	-0.049	-0.057*	-0.104***	-0.098***	-0.098***	-0.069***
	(0.036)	(0.031)	(0.027)	(0.028)	(0.028)	(0.017)
Father Sec Edu				0.075**	0.061*	0.105**
				(0.036)	(0.036)	(0.049)
Father Ter Edu				0.180***	0.134**	0.173***
				(0.048)	(0.054)	(0.066)
Mother Sec Edu				0.013	-0.002	0.042
				(0.037)	(0.037)	(0.066)
Mother Ter Edu				0.060	0.013	0.036
				(0.047)	(0.048)	(0.089)
Mother Working $\times$ ISEI					0.003***	0.003***
					(0.001)	(0.001)
Father Working $\times$ ISEI					0.003***	0.004***
					(0.001)	(0.001)
N	34161	34161	34161	34161	34161	21385
# Country *m*	48	48	48	48	48	41
R Squared	0.14	0.25	0.66	0.67	0.67	0.63
Host Country FE	No	Yes	Yes	Yes	Yes	Yes
School FE	No	No	Yes	Yes	Yes	Yes

Notes: The Table shows results for second generation immigrants on the mother's side. The sample includes only cases where both parents report a country of origin and the country of origin of the mother runs a PISA test on natives. *Score Science Country m* is the average science PISA score of natives (standardized to have mean 0 and standard deviation 1 across all countries participating to the test) in the country of birth of the mother, across all available waves. All specifications control for intercept, students' age (in months), wave fixed effect and a dummy for father's immigrant status. Observations are weighted according to the provided sample weights. Standard errors are clustered by mother's country of origin, and inflated by the estimated measurement error in test scores. * denotes significance at 10%, ** at 5%, *** at 1%.

A.2.1.4 Standard Errors

Throughout the paper, standard errors for the analyses on PISA data are constructed taking into account the fact that student performance is reported through plausible values. Using the average of the five plausible values as a measure of individual performance guarantees unbiased estimates of group-level means and regression coefficients; however, measures of dispersion need to take into account the within-student variability in plausible values.

As recommended in OECD (2009), for the purpose of computing standard errors all regression with individual test scores as dependent variable are estimated five times, using all plausible values in turn. For each regression we employ an estimator for the sampling variance clustered at the level of the mother's country of origin. The final sampling variance, SV, is given by the average of the sampling variances obtained with the five plausible values.

In addition, standard errors are inflated by the imputation variance due to the fact that test scores measure the latent student's skills with error. The imputation variance, IV, is estimated as the average squared deviation between the estimates obtained with each plausible value and the final estimate (obtained using the average of the plausible values), with the appropriate degree of freedom adjustment.

Finally, as shown in Little and Rubin (1987), the final error variance TV can be obtained by combining the sampling and imputation variance in

$$TV = SV + \left(1 + \frac{1}{K}\right) IV$$

where $K = 5$ is the number of plausible values for each student. The final standard errors are given by the squared roots of the final error variances.

As an alternative to estimate SV, OECD (2009) recommends to apply Fay's variant of the Balanced Repeated Replication (BRR) method, which directly takes into account the two-stage stratified sampling design of the PISA test. This is implemented by iterating each regression over the 80 sets of replicate weights provided in the PISA dataset. The sampling variance estimate is then given by the average squared deviation between the replicated estimates and the estimate obtained with final weights, with a degree of freedom correction depending on the Fay coefficient (a parameter that governs the variability between different sets of replicate weights).

Table A.8 shows the resulting standard errors for our baseline specification. For computational convenience, we implemented the "unbiased shortcut" procedure described in OECD (2009), which uses only one set of plausible values to estimate the sampling variance (while the imputation variance is estimated using all five sets, as described above). In all specifications, the standard error on our coefficient of interest is smaller compared to Table 1.2 in the main text, suggesting that our clustered sampling variance is rather conservative.

Table A.8: Main results-PISA (BRR Standard Errors)

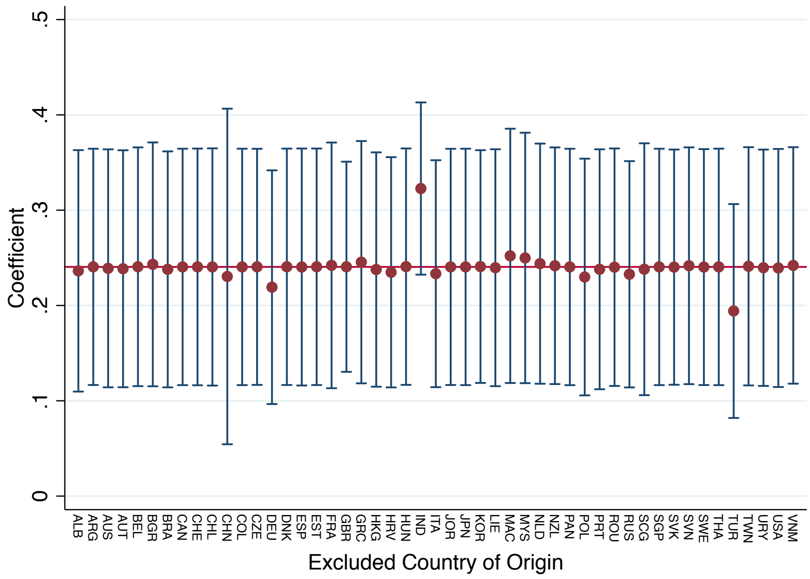
	Dependent Variable: Math Test Score					
	[1]	[2]	[3]	[4]	[5]	[6]
			All			No East Asia
Score Country *m*	0.662***	0.499***	0.253***	0.249***	0.240***	0.225***
	(0.024)	(0.046)	(0.034)	(0.034)	(0.034)	(0.053)
Female	-0.140***	-0.148***	-0.206***	-0.204***	-0.201***	-0.187***
	(0.025)	(0.022)	(0.019)	(0.018)	(0.018)	(0.023)
Father Sec Edu				0.030	0.014	0.022
				(0.030)	(0.030)	(0.056)
Father Ter Edu				0.099**	0.045	0.049
				(0.039)	(0.042)	(0.063)
Mother Sec Edu				0.001	-0.015	0.027
				(0.027)	(0.026)	(0.046)
Mother Ter Edu				0.032	-0.011	0.023
				(0.034)	(0.034)	(0.050)
Mother Working $\times$ ISEI					0.003***	0.003***
					(0.001)	(0.001)
Father Working $\times$ ISEI					0.003***	0.003***
					(0.001)	(0.001)
N	40067	40067	40067	40067	40067	25454
# Country m	49	49	49	49	49	42
R Squared	0.16	0.25	0.67	0.67	0.67	0.63
Host Country FE	No	Yes	Yes	Yes	Yes	Yes
School FE	No	No	Yes	Yes	Yes	Yes

Notes: The Table shows results for second generation immigrants on the mother's side. The sample and specifications are the same as in Table 1.2 in the main text. Standard errors are computed using the provided replicate weights, and inflated by the estimated measurement error in test scores. The sampling variance is estimated through the "unbiased shortcut" procedure described in OECD (2009). * denotes significance at 10%, ** at 5%, *** at 1%.

A.2.1.5 Excluding Single Countries

In this section we investigate to what extent our results are driven by specific countries of origin or host countries. Figure A.1 shows the estimated coefficient of interest when countries of origin are excluded one by one. The resulting estimates are never significantly different from the baseline, represented by the horizontal line. Even if the difference is insignificant, the coefficient is substantially higher when second generation students from India are excluded; this reflect the fact that these students are outliers since they perform relatively well even though, across natives, India is near the bottom of the international ranking. On the other had, the coefficient becomes somewhat smaller when second generation immigrants from Germany, Poland and Turkey are excluded. Overall, the statistical significance and the rough magnitude of our coefficient of interest is not driven by any specific country of origin.

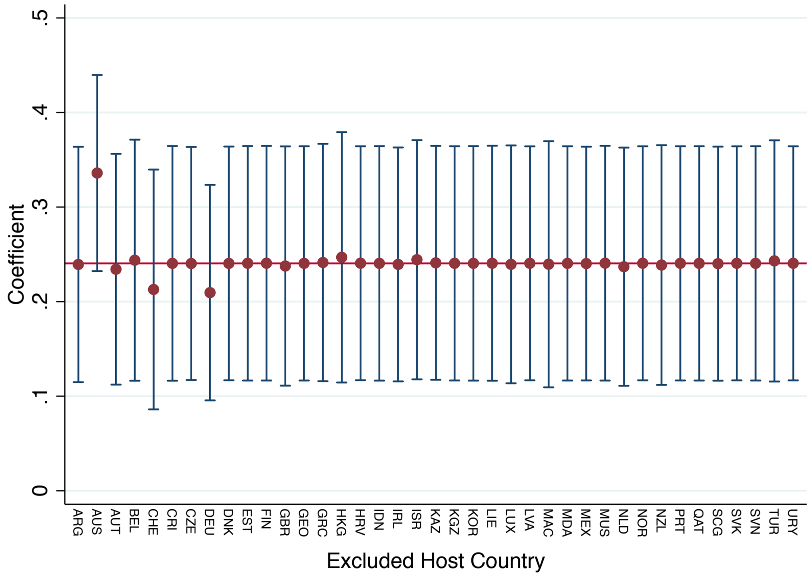
Figure A.1: Reduced Form Coefficient when Excluding Countries of Origin One by One



Notes: The Figure plots the estimated coefficients and 95% confidence intervals on the average PISA score of natives in mother's country of origin, with the dependent variable and other controls being the same as in column 5 of Table 1.2. Each dot corresponds to a different specification, where students with mothers from the indicated country of origin are excluded. Standard errors are clustered by mother's country of origin.

Figure A.2 shows the result from the corresponding exercise on host countries. The coefficient is positive, significant and quite stable across the 39 specifications. As a partial exception, the coefficient is quite a bit higher (even though the difference is not statistically significant) when second generation immigrants in Australia are excluded from the sample. While in principle this might be due to a number of factors, a possible rationalization is the relatively stronger negative selection of East Asian emigrant parents to Australia, given the geographic proximity.

Figure A.2: Reduced Form Coefficient when Excluding Host Countries One by One



Notes: The Figure plots the estimated coefficients and 95% confidence intervals on the average PISA score of natives in mother's country of origin, with the dependent variable and other controls being the same as in column 5 of Table 1.2. Each dot corresponds to a different specification, where students in the indicated host country are excluded. Standard errors are clustered by mother's country of origin.

A.2.1.6 Alternative Measures of School Performance

Table A.9 shows results from our baseline specification with different measures of school performance available in the PISA data as outcome variables. Columns 1 and 2 consider the grade attended by each student, whose variation within country and within school (once age is controlled for) will be driven by either grade repetition or grade skipping. Columns 3 and 4 consider a measure of truancy, that is the number of school days each student has skipped in the previous two weeks; this measure is available for the 2012 wave only. According to both measures, second generation immigrants from high PISA countries perform better. Within the same school, an extra individual-level standard deviation in the average performance of native students in the mother's country of origin corresponds to 0.07 extra grades (9% of a standard deviation) and 0.041 less school days skipped (13% of a standard deviation). The gap is marginally larger if we do not include school fixed effects.

Table A.9: Alternative Measures of School Performance

	Grade		Truancy	
	[1]	[2]	[3]	[4]
Score Country *m*	0.085*	0.069*	-0.046***	-0.041*
	(0.050)	(0.039)	(0.017)	(0.025)
Female	0.119***	0.092***	-0.014	-0.010
	(0.020)	(0.015)	(0.019)	(0.015)
Father Sec Edu	0.049*	0.017	-0.050	-0.018
	(0.026)	(0.040)	(0.038)	(0.039)
Father Ter Edu	0.047	0.016	-0.045	-0.022
	(0.033)	(0.042)	(0.045)	(0.039)
Mother Sec Edu	0.067**	0.000	-0.004	0.008
	(0.033)	(0.021)	(0.039)	(0.039)
Mother Ter Edu	0.096**	-0.003	-0.007	-0.005
	(0.043)	(0.030)	(0.037)	(0.031)
Mother Working × ISEI	0.002***	-0.000	0.000	-0.000
	(0.000)	(0.001)	(0.000)	(0.000)
Father Working × ISEI	0.001*	0.000	-0.000	-0.000
	(0.001)	(0.001)	(0.000)	(0.000)
N	40017	40017	12368	12368
# Country *m*	49	49	46	46
Mean Dep Var	9.38	9.38	0.11	0.11
SD Dep Var	0.75	0.75	0.31	0.31
R Squared	0.44	0.78	0.18	0.60
Host Country FE	Yes	Yes	Yes	Yes
School FE	No	Yes	No	Yes

Notes: The Table shows results for second generation immigrants on the mother's side. The dependent variable is grade attended in columns 1 and 2 and the number of school days skipped in the previous two weeks in columns 3 and 4. The controls are the same as in column 5 of Table 1.2 in the main text. Standard errors are clustered by mother's country of origin. * denotes significance at 10%, ** at 5%, *** at 1%.

A.2.1.7 Alternative Measures of Socio-Economic Status

Table A.10 considers alternative measures of parental socio-economic status available from the PISA questionnaires. In Column 2 we control for an index of family wealth, based on the presence and the number of various items in students' homes, including computers, cars, cellular phones, televisions and rooms with bath or shower. Column 3 includes an index of home possessions, which is based on all elements in the wealth index and additionally considers books, various educational resources and pieces of classical culture. Column 4 considers the broadest measure available in PISA, an index of Economic, Social and Cultural Status (ESCS) which combines home possessions with information on parents' education and occupational status. All indexes are standardized to take mean 0 and (individual-level) standard deviation 1 across the 73 countries (pooled, equally weighted) participating to the test.

The results are very similar compared to the baseline specification, reported in column 1. The magnitude of our coefficient of interest varies little across specifications, even when (in Column 5) we introduce all indexes of socio-economic status in the same regression. Home possessions and the ESCS index are positively related to students' performance, while wealth is not.[1] Overall, the results suggest the controlling further for observable measures of socio-

[1]Much of the variation in wealth seems to be absorbed by the school fixed effect, since this index enters

economic background does not affect affect the magnitude of our estimated parental component.

Table A.10: Alternative Measures of Socio-economic Status

	Dependent Variable: Math Test Score				
	[1]	[2]	[3]	[4]	[5]
Score Country *m*	0.240***	0.272***	0.248***	0.236***	0.261***
	(0.065)	(0.071)	(0.068)	(0.065)	(0.068)
Female	-0.201***	-0.190***	-0.206***	-0.204***	-0.198***
	(0.022)	(0.027)	(0.022)	(0.023)	(0.026)
Father Sec Edu	0.014	0.045*	0.018		
	(0.022)	(0.026)	(0.021)		
Father Ter Edu	0.045	0.112***	0.076**		
	(0.034)	(0.037)	(0.031)		
Mother Sec Edu	-0.015	-0.005	-0.007		
	(0.037)	(0.039)	(0.039)		
Mother Ter Edu	-0.011	0.042	0.008		
	(0.042)	(0.044)	(0.041)		
Mother Working $\times$ ISEI	0.003***				
	(0.001)				
Father Working $\times$ ISEI	0.003***				
	(0.001)				
Wealth		-0.004			-0.211***
		(0.023)			(0.035)
Home Possessions			0.087***		0.180***
			(0.019)		(0.039)
ESCS				0.111***	0.096***
				(0.029)	(0.032)
N	40067	34134	40049	40060	34134
# Country m	49	48	49	49	48
R Squared	0.67	0.68	0.67	0.67	0.69
Host Country FE	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes

Notes: The Table shows results for second generation immigrants on the mother's side. The sample includes only cases where both parents report a country of origin and the country of origin of the mother runs a PISA test on natives. *Score Country m* refers to the average math PISA score of natives (standardized to have mean 0 and standard deviation 1 across the OECD countries participating in 2003) in the country of birth of the mother, across all available waves. All specifications control for intercept, students' exact age (in months), wave fixed effect and a dummy for father immigrant status; specification 1 additionally controls for dummies for parents' employment status (full-time employed, part-time employed, not working). *Working* refers to either full-time or part-time employed. *Wealth*, *Home Possessions* and *ESCS* are indexes of socio-economic status, discussed in the text. Observations weighted according to the provided sample weights. Standard errors are clustered by mother's country of origin, and inflated by the estimated measurement error in test scores. * denotes significance at 10%, ** denotes significance at 5%, *** denotes significance at 1%.

positively and significantly in a specification with host country fixed effects (results not shown, available upon request).

A.2.2 US Census

A.2.2.1 Results for Second Generation Immigrants on the Father's Side

Table A.11: Main Results for Fathers - US CENSUS

	Dependent variable: 1 = Never repeated a grade					
	[1]	[2]	[3]	[4]	[5]	[6]
			All			No East Asia
Score Country *f*	0.102***	0.063***	0.038***	0.033***	0.029***	0.023*
	(0.037)	(0.020)	(0.012)	(0.011)	(0.011)	(0.013)
Female	0.072***	0.071***	0.072***	0.072***	0.072***	0.073***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Mother Sec Edu			0.066***	0.065***	0.059***	0.060***
			(0.017)	(0.017)	(0.015)	(0.016)
Mother Ter Edu			0.086***	0.085***	0.071***	0.073***
			(0.013)	(0.014)	(0.012)	(0.013)
Father Sec Edu			0.035***	0.031***	0.027***	0.029***
			(0.007)	(0.007)	(0.007)	(0.007)
Father Ter Edu			0.057***	0.053***	0.041***	0.044***
			(0.009)	(0.009)	(0.007)	(0.007)
Log Family Income					0.034***	0.035***
					(0.005)	(0.005)
N	46410	46410	46410	46410	46410	43909
# Country f	61	61	61	61	61	54
R Squared	0.07	0.10	0.11	0.12	0.12	0.12
Comm Zone FE	No	Yes	Yes	Yes	Yes	Yes
Years Since Migr Father	No	No	No	Yes	Yes	Yes

Notes: The Table shows results for second generation immigrants on the father's side. *Score Country f* is the average math PISA score of natives (standardized to have mean 0 and standard deviation 1 across all countries participating to the test) in the country of birth of the father, across all available waves. All specifications control for intercept, child age dummies, parents' age, number of siblings, year fixed effect, (year-specific) quarter of birth fixed effect and mother's immigrant status. Observations weighted according to the provided sample weights. Standard errors are clustered by father's country of origin. * denotes significance at 10%, ** at 5%, *** at 1%.

A.2.2.2 Results for Second Generation Immigrants and Natives

Table A.12: Main Results for Mothers and Fathers - US CENSUS

	Dependent variable: 1 = Never repeated a grade					
	[1]	[2]	[3]	[4]	[5]	[6]
			All			No East Asia
Score Country *m*	0.050**	0.036***	0.018*	0.015*	0.013*	0.010
	(0.022)	(0.013)	(0.009)	(0.008)	(0.008)	(0.009)
Score Country *f*	0.075**	0.057***	0.032***	0.026***	0.023***	0.020*
	(0.030)	(0.020)	(0.010)	(0.009)	(0.008)	(0.010)
Native Mother	-0.005	-0.005	-0.071***	0.000	-0.071***	-0.071***
	(0.007)	(0.005)	(0.016)	(0.000)	(0.018)	(0.019)
Native Father	-0.002	0.002	-0.059***	-0.066***	-0.061***	-0.058***
	(0.009)	(0.006)	(0.015)	(0.016)	(0.015)	(0.015)
Female	0.084***	0.084***	0.085***	0.085***	0.085***	0.085***
	(0.001)	(0.000)	(0.000)	(0.001)	(0.001)	(0.000)
Mother Sec Edu			0.060***	0.058***	0.055***	0.055***
			(0.015)	(0.015)	(0.015)	(0.016)
Mother Ter Edu			0.077***	0.075***	0.065***	0.062***
			(0.015)	(0.015)	(0.015)	(0.016)
Father Sec Edu			0.040***	0.038***	0.033***	0.035***
			(0.010)	(0.010)	(0.010)	(0.009)
Father Ter Edu			0.061***	0.059***	0.045***	0.048***
			(0.014)	(0.013)	(0.013)	(0.013)
Native Mother × Mother Sec Edu			0.059***	0.061***	0.055***	0.055***
			(0.015)	(0.015)	(0.015)	(0.016)
Native Mother × Mother Ter Edu			0.068***	0.069***	0.063***	0.066***
			(0.015)	(0.015)	(0.015)	(0.016)
Native Father × Father Sec Edu			0.043***	0.045***	0.041***	0.039***
			(0.010)	(0.010)	(0.010)	(0.009)
Native Father × Father Ter Edu			0.058***	0.059***	0.056***	0.052***
			(0.014)	(0.014)	(0.013)	(0.013)
Log Family Income					0.035***	0.035***
					(0.000)	(0.000)
N	1299079	1299079	1299079	1299079	1292410	1288059
# Country *m*	61	61	61	61	61	54
# Country *f*	61	61	61	61	61	54
R Squared	0.04	0.05	0.07	0.07	0.08	0.08
Comm Zone FE	No	Yes	Yes	Yes	Yes	Yes
Years Since Migr Mother	No	No	No	Yes	Yes	Yes
Years Since Migr Father	No	No	No	Yes	Yes	Yes

Notes: The Table shows results for second generation immigrants and natives. Sample includes only cases where both parents report a country of origin that runs a PISA test on natives. *Score Country m* and *Score Country f* are the average math PISA score of natives (standardized to have mean 0 and standard deviation 1) in the country of birth of the mother and father, across all available waves. All specifications control for intercept, child age dummies, parents' age, number of siblings, log family income, year fixed effect and (year-specific) quarter of birth fixed effect. Observations are weighted accordind to the provided sample weights. Robust standard errors clustered by mother's and father's country of origin. * denotes significance at 10%, ** denotes significance at 5%, *** denotes significance at 1%.

A.3 Additional Results on Selection

A.3.1 Selection into Schools

Table A.13 shows regressions with several proxies for school quality (constructed from the information available in the School Questionnaire) as outcome variables. After controlling for country fixed effects and the usual observable characteristics, a higher PISA score in the country of origin of the mother is associated with schools where natives score better in the PISA test, no matter whether we take the raw average (column 1) or clean it from observable characteristics (column 2), where admissions are more likely to be based on academic records, the proportion of teachers with at least some tertiary education is higher and the proportion of students dropping out is lower.

Table A.13: Selection into Schools

	Avg Score School	Estimated School FE	Academic Admission	Share Qual Teachers	Dropout Rate
	[1]	[2]	[3]	[4]	[5]
Score Country *m*	0.167*	0.154**	0.045*	0.026**	-0.016**
	(0.087)	(0.076)	(0.026)	(0.012)	(0.007)
Female	0.046**	0.051***	0.009	0.004	-0.003
	(0.019)	(0.018)	(0.009)	(0.007)	(0.003)
Father Sec Edu	0.074**	0.061*	0.053*	-0.021	0.005
	(0.029)	(0.034)	(0.029)	(0.014)	(0.010)
Father Ter Edu	0.148***	0.127***	0.072**	-0.005	-0.001
	(0.037)	(0.038)	(0.028)	(0.017)	(0.011)
Mother Sec Edu	0.114***	0.100***	0.015	-0.004	-0.034
	(0.024)	(0.022)	(0.018)	(0.010)	(0.028)
Mother Ter Edu	0.146***	0.123***	0.009	-0.010	-0.037
	(0.036)	(0.037)	(0.027)	(0.016)	(0.026)
Mother Working × ISEI	0.005***	0.004***	0.001***	0.001***	-0.000
	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)
Father Working × ISEI	0.005***	0.004***	0.000	0.001***	-0.000**
	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
N	39573	39448	40067	29893	9751
# Country *m*	49	49	49	48	41
Mean Dep. Var.	0.46	0.50	0.75	0.80	0.03
St. Dev. Dep. Var.	0.59	0.56	0.43	0.28	0.07
R Squared	0.37	0.39	0.18	0.42	0.17
Host Country FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes

Notes: The sample includes only second generation immigrants on the mother's side. *Avg Score School* is the average math PISA score of native students in the same school, *Estimated School FE* is the estimated school fixed effect in a regression of the math PISA score on gender, age in months, parental education and occupational status (limiting the sample to native students), *Academic Admission* is a dummy that takes value 1 whenever schools report that student's record of academic performance is either *always* or *sometimes* considered for admissions, *Share Qual Teachers* is the share of current teachers with at least the ISCED 5A level of education and *Dropout Rate* is the share of students who leave the school without having obtained the corresponding diploma. *Score Country m* is the average math PISA score of natives (standardized to have mean 0 and standard deviation 1 across all countries participating to the test) in the country of birth of the mother, across all available waves. All specifications control for intercept, students' age (in months), wave fixed effect, a dummy for father's immigrant status and dummies for parents' employment status (full-time employed, part-time employed, not working). *Working* refers to either full-time or part-time employed. Standard errors are clustered by mother's country of origin. * denotes significance at 10%, ** at 5%, *** at 1%.

A.3.2 Selection into Host Countries

An additional concern is that immigrant parents from high PISA countries may be systematically selecting host countries (or schools) where, because of idiosyncratic factors, it is easier for them and their children to integrate and perform well. Of course the quality of the match between parents or children on one side and countries or schools on the other is unobservable, and it is difficult to rule out entirely this possibility. However, we can check whether immigrants from high PISA countries are located in countries which are, according to reasonable proxies, culturally closer to their country of origin. Table A.14 explores this possibility. In column 2 we add to the baseline regression of column 1 a dummy variable that takes value 1 for all students that declare to speak a foreign language at home (which is available only for part of the sample). While the coefficient on this newly added control is, as expected, negative and significant, our main coefficient of interest is virtually unaffected.

In column 4 and 6 we add to the baseline specifications (reported in columns 3 and 5 respectively) controls for linguistic distance (constructed through the softwares provided by the Automated Similarity Judgment Program (Wichmann and Brown, 2016)) and cultural distance (from Spolaore and Wacziarg (2015)); both measures are standardized to have mean 0 and standard deviation 1 across all country pairs in the sample. In both cases the impact on our coefficient of interest is positive and of negligible magnitude.[2]

[2] In recent work, Isphording et al. (2016) argue that linguistic distance impacts immigrant students' mathematics performance through its effect on reading skills. These results are not in contrast with ours given that we are looking at linguistic distance for immigrant parents, while all students in our sample are born in the country where they attend school.

Table A.14: Linguistic and Cultural Distance

	Dependent Variable: Math Test Score					
	[1]	[2]	[3]	[4]	[5]	[6]
Score Country *m*	0.220***	0.217***	0.239***	0.241***	0.216**	0.241***
	(0.062)	(0.060)	(0.067)	(0.065)	(0.088)	(0.088)
Female	-0.192***	-0.194***	-0.201***	-0.200***	-0.202***	-0.201***
	(0.026)	(0.027)	(0.022)	(0.022)	(0.023)	(0.023)
Father Sec Edu	0.012	0.010	0.015	0.015	-0.015	-0.018
	(0.022)	(0.023)	(0.021)	(0.021)	(0.057)	(0.059)
Father Ter Edu	0.034	0.032	0.041	0.042	0.014	0.009
	(0.035)	(0.035)	(0.031)	(0.030)	(0.069)	(0.072)
Mother Sec Edu	-0.030	-0.035	-0.013	-0.012	0.007	0.012
	(0.033)	(0.033)	(0.039)	(0.039)	(0.101)	(0.108)
Mother Ter Edu	-0.012	-0.017	-0.012	-0.010	-0.050	-0.044
	(0.038)	(0.039)	(0.043)	(0.044)	(0.122)	(0.130)
Mother Working × Mother ISEI	0.002***	0.002***	0.003***	0.003***	0.003**	0.003**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)
Father Working × Father ISEI	0.003***	0.003***	0.003***	0.003***	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)
Foreign Language at Home		-0.063**				
		(0.028)				
Mother Linguistic Distance				-0.001		
				(0.013)		
Father Linguistic Distance				0.012		
				(0.009)		
Mother Cultural Distance						0.070
						(0.075)
Father Cultural Distance						-0.053
						(0.082)
N	37827	37827	38487	38487	10309	10309
# Country m	49	49	49	49	35	35
R Squared	0.67	0.67	0.67	0.67	0.68	0.69
Host Country FE	Yes	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The Table shows results for second generation immigrants on the mother's side, augmented for controls for linguistic and cultural distance. Sample includes only cases where both parents report a country of origin and the country of origin of the mother participates to PISA. *Score Country m* is the average math PISA score of natives (standardized to have mean 0 and standard deviation 1 across all countries participating to the test) in the country of birth of the mother, across all available waves. All specifications control for intercept, students' age (in months), wave fixed effect, a dummy for father's immigrant status and dummies for parents' employment status (full-time employed, part-time employed, not working). *Working* refers to either full-time or part-time employed. *Linguistic Distance* and *Cultural Distance* vary at the country-pair level, and are standardized to take mean 0 and standard deviation 1 across all country pairs in the sample (sources are discussed in the main text). Observations are weighted according to the provided sample weights. Standard errors are clustered by mother's country of origin, and inflated by the estimated measurement error in test scores. * denotes significance at 10%, ** at 5%, *** at 1%.

A.3.3 Insights from the Migration Literature

The migration literature has extensively debated the country-level determinants of emigrants' self-selection in terms of observable and unobservable skills. While, to our knowledge, the PISA score itself has not been explicitly considered in this literature, this variable is correlated with several others that have been advocated as measuring direct determinants of selection. In Figure A.3 we plot some of these variables against the PISA score of native students in the country of origin; since in the PISA sample we do not know the exact date of migration, we use data on selection determinants in 1985 or the closest available data, which should plausibly approximate the pre-migration conditions for the average migrant in our sample.[3]

First, the seminal contribution of Borjas (1987) gives a central role to the difference in income inequality between the origin and destination countries, predicting positive selection if the wage structure of the host country is such that skills are rewarded more compared to the country of origin, and negative selection in the opposite case. Panels (a) and (b) of Figure A.3 show that on average emigrant parents from high PISA countries do emigrate to countries more unequal (as measured by the Gini coefficient and the estimated return to education) than their countries of origin, implying that they would be more positively selected according to Borjas' theory.[4] However, this theory has received mixed support (Chiquiar and Hanson, 2005; Belot and Hatton, 2012), and in particular Grogger and Hanson (2011) argue that the absolute (as opposed to the relative) wage gap between high and low earners provides a better rationalization of the patterns of selection observed in the data. Panel (c) shows that, according to the preferred measure in Grogger and Hanson (2011), emigrants from high PISA countries (if anything) face a relatively lower absolute earning spread in their host countries, implying that they would be more negatively selected.[5]

Another strand of the literature emphasizes the importance of liquidity constraints (Chiswick, 2000; Belot and Hatton, 2012). These papers suggest that emigrants' self-selection should be more negative from richer countries, where facing emigration costs is affordable for a larger share of the population. Since the average PISA score is positively correlated with real GDP in 1985 (Panel d), we should expect negative differential selection according to this mechanism as well. Panel (e) shows instead the extent to which emigrants choose countries with a large pre-existing community from the same country of origin, since McKenzie and Rapoport (2010), among others, argue that stronger social networks act to reduce the effective cost of migration inducing negative selection.[6] China is an outlier in this dimension, since many Chinese parents are observed in Macao and Hong Kong, where Chinese-born represented respectively the 37% and 36% of the population in 1980; therefore, this "chain migration" view would predict negative selection for China, and no systematic pattern of differential selection for the other countries.[7] Finally, Panel (f) shows that emigrants from high PISA countries are not systemat-

[3]In the US Census, where we observe years since migration, the average mother of a US-born 15-year-old student migrated 20 years earlier.

[4]We take the Gini Index from the cross-country dataset constructed in Brueckner and Lederman (2015), and we use the 1985 observation when available and 1990 or 1995 when not. The Mincerian coefficients come from Psacharopoulos and Patrinos (2004), who collect estimates from a large set of papers; most observations refer to the 1980s.

[5]Grogger and Hanson (2011) combine information from the Luxembourg Income Study and the WIDER dataset to construct an estimate of the absolute income gap (in thousands of 2000 US dollars) between the 80^{th} and 20^{th} percentiles of the income distribution in each country.

[6]We construct a matrix of bilateral migration shares in 1980 from the Global Bilateral Migration Database, discussed in Ozden et al. (2011). Each entry of this matrix gives us the share of the resident population in country i that was born in country j.

[7]The results of the paper are robust to the omission of Macao and Hong Kong as host countries, and to their aggregation to China as well. If anything, the relative over-performance of Chinese second-generation immigrants compared to other countries of origin is weaker in these two countries, perhaps due to the patterns of selection

ically located in a country closer or farther from their country of origin.[8] This is relevant since geographical distance has been shown to be associated with negative selection (Grogger and Hanson, 2011; Belot and Hatton, 2012), most likely due to its effect on the cost of migrating.

Recent work by Albornoz et al. (2012) examines theoretically the determinants of selection in terms of parental motivation for their children's education, which might be only partially correlated with parents' skills. Among other channels, the authors stress the importance of the relative quality of the school systems in the host and source countries, since highly motivated parents are more likely to migrate to countries with better educational prospects for their children.[9] Under the presumption that high PISA countries have better schools on average, parents emigrating from these countries should be, ceteris paribus, relatively more negatively selected.

All in all, given the determinants of self-selection considered in the literature, we conclude that a pattern of (weakly) negative differential selection should be expected.

A.3.4 Selection Analysis for the Census Data

In this Appendix we provide a discussion of the patterns of differential selection in the US Census data. While the analysis parallels the one in the main text on the PISA sample, the information on years since migration available in the Census allows us to implement additional checks.

In order to benchmark emigrant parents against non-emigrants in their country of origin, we use school attainment data from Barro and Lee (2013), combined with information on the duration of primary and secondary school in each country from the World Development Indicators, to construct estimates for the average and the standard deviation of years of education in the across countries of origin.[10] Differently from the PISA data, we cannot build these measures for parents of school-age children only; we can, however, restrict attention to adults between 35 and 45 years of age. At the individual level, our proxy for selection is therefore years of education standardized by the (gender-specific) average and standard deviation in the country of origin. At the country level, we simply take the average of this measure.

In Figure A.4 we plot these country of origin-level averages against the PISA score of native students in those countries. Similarly to the PISA sample, we find a weakly negative pattern, suggesting that parents from high PISA countries are somewhat more negatively selected. In Table A.15 we check whether this pattern arises also when we include commuting zone fixed effects: for both mothers and fathers, the coefficients are negative and not statistically different from 0.

discussed in this section.

[8]The geographical distance data comes from the CEPII's GeoDist dataset (Mayer and Zignago, 2011). We use the simple distance between the most populated cities, expressed in kilometers.

[9]Other determinants of selection considered in Albornoz et al. (2012) are the absolute skill premia in host and source countries and migration costs. As discussed above, the available evidence on these dimensions suggests that, if anything, we should expect parents emigrating from high PISA countries to be relatively negatively selected.

[10]Following Barro and Lee (2013), we impute a duration of 4 years for tertiary education in all countries.

Table A.15: Selection - US Census

	Dependent Variable: Standardized Years of Education	
	[1] Mothers	[2] Fathers
Score Country *m*	-0.282	
	(0.478)	
Score Country *f*		-0.009
		(0.393)
N	52875	46254
R Squared	0.06	0.08
Year FE	Yes	Yes
Comm Zone FE	Yes	Yes

Notes: The Table shows results for emigrant mothers in specification (1) and emigrant fathers in specification (2). The dependent variable is years of education standardized by the average and standard deviation of mothers' (specification 1) and fathers' (specification 2) years of education in the country of origin. *Score Country m* and *Score Country f* are the average math PISA scores of natives (standardized to have mean 0 and standard deviation 1 across all countries participating to the test) in the country of birth of the mother and the father, across all available waves. All specifications control for intercept and wave fixed effect. Standard errors are clustered by mother's country of origin in specification (1) and by father's country of origin in specifications (2). * denotes significance at 10%, ** at 5%, *** at 1%.

One of the concerns highlighted in Section 3.4 was that years of education are not necessarily pre-determined with respect to migration, and parents might have acquired more or less education as a consequence of their migration decision (and, importantly, for our purposes, might have done so differentially from different countries of origin). We can make some progress in testing this hypothesis by analyzig selection patterns for parents that completed their education in their home country, since for those individuals the relative quality of the US school system should have not played any role in their education choices (and therefore education is more likely to represent a good proxy of pre-determined skills). Figure A.5 and Table A.16 are the counterparts of Figure A.4 and Table A.15 when the sample is restricted only to parents more likely to have completed their education before migrating to the US (see Section 1.7 for a description of how these parents are identified based on the available information). For both mothers and fathers, the pattern of differential selection is weakly negative with respect to the average PISA score, and not very different from the one obtained in the full sample.

Table A.16: Selection - US Census (Parents Entirely Educated in Home Country)

	Dependent Variable: Standardized Years of Education	
	[1]	[2]
	Mothers	Fathers
Score Country *m*	-0.182	
	(0.357)	
Score Country *f*		-0.111
		(0.299)
N	29851	27070
R Squared	0.07	0.09
Year FE	Yes	Yes
Comm Zone FE	Yes	Yes

Notes: The Table shows results for emigrant mothers in specification (1) and emigrant fathers in specification (2). In all specifications, the sample includes only cases where the parent was entirely educated in his or her home country. The dependent variable is years of education standardized by the average and standard deviation of mothers' (specification 1) and fathers' (specification 2) years of education in the country of origin. *Score Country m* and *Score Country f* are the average math PISA score of natives (standardized to have mean 0 and standard deviation 1 across all countries participating to the test) in the country of birth of the mother and the father, across all available waves. All specifications control for intercept and wave fixed effect. Standard errors are clustered by mother's country of origin in specification (1) and by father's country of origin in specification (2). * denotes significance at 10%, ** at 5%, *** at 1%.

Still, parents might have based their educational choices based on their future relocation to the US, and perhaps this might bias the pattern of selection differentially across countries. It would be worrying if differential selection turned out to be negative only for those parents for whom migration is likely to have played a bigger role in their educational choices, and perhaps positive for the rest of the sample. To check for this possibility, in Table A.17 we present results from specifications where we interact the average PISA score in the country of origin with the number of years between education completion and migration (still restricting the sample to parents entirely educated in their home country). The underlying idea is that the more time has passed between education completion and migration, the less is likely that educational choices were made taking future relocation into account, and the closer we get to the ideal situation where education truly reflects skills pre-determined with respect to migration. For both mothers and fathers, the coefficient on the interaction term is positive but not statistically significant, and its magnitude is so small that the pattern of differential selection would not be positive and significant for any gap between education completion and migration observed in the sample. This result gives us some further confidence that our findings on selection are not driven by a differential effect of migration on parental education.

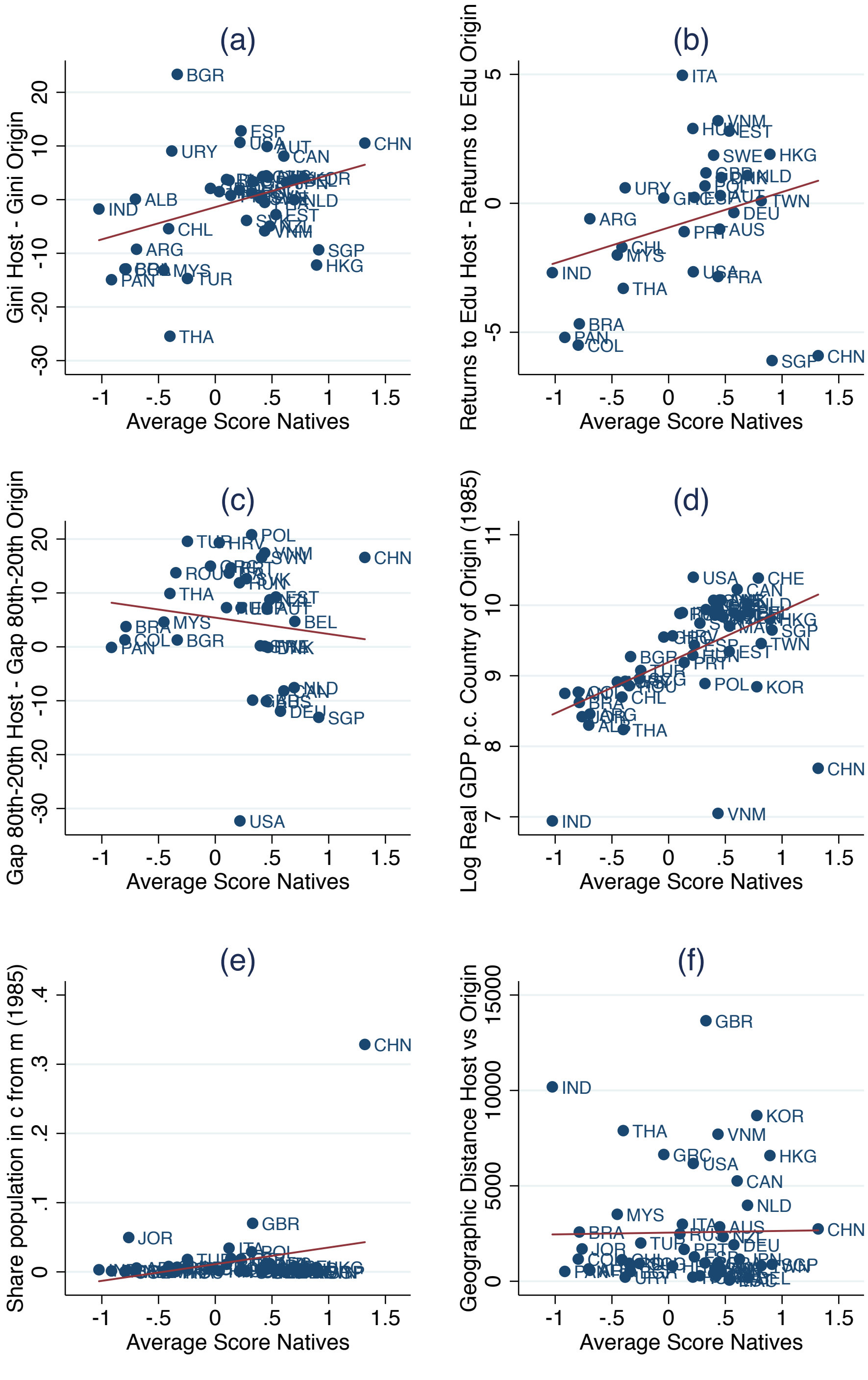
Table A.17: Selection - Heterogeneity with respect to Education Completion and Migration Dates

	Dependent Variable: Standardized Years of Education	
	[1] Mothers	[2] Fathers
Score Country *m*	-0.330	
	(0.497)	
Score Country *m* × Years betw Edu and Migration Mother	0.006	
	(0.018)	
Years betw Edu and Migration Mother	-0.056***	
	(0.006)	
Score Country *f*		-0.248
		(0.381)
Score Country *f* × Years betw Edu and Migration Father		0.004
		(0.012)
Years betw Edu and Migration Father		-0.053***
		(0.005)
N	29851	27070
R Squared	0.16	0.22
Year FE	Yes	Yes
Comm Zone FE	Yes	Yes

Notes: The Table shows results for emigrant mothers in specification (1) and emigrant fathers in specification (2). In all specifications, the sample includes only cases where the parent was entirely educated in his or her home country. The dependent variable is years of education standardized by the average and standard deviation of mothers' (specification 1) and fathers' (specification 2) years of education in the country of origin. *Score Country m* and *Score Country f* are the average math PISA score of natives (standardized to have mean 0 and standard deviation 1 across all countries participating to the test) in the country of birth of the mother and the father, across all available waves. *Years betw Edu and Migration* refers to the number of years occurred between education completion (imputed from the educational attainment) and migration to the US. All specifications control for intercept and wave fixed effect. Standard errors are clustered by mother's country of origin in specification (1) and by father's country of origin in specification (2). * denotes significance at 10%, ** at 5%, *** at 1%.

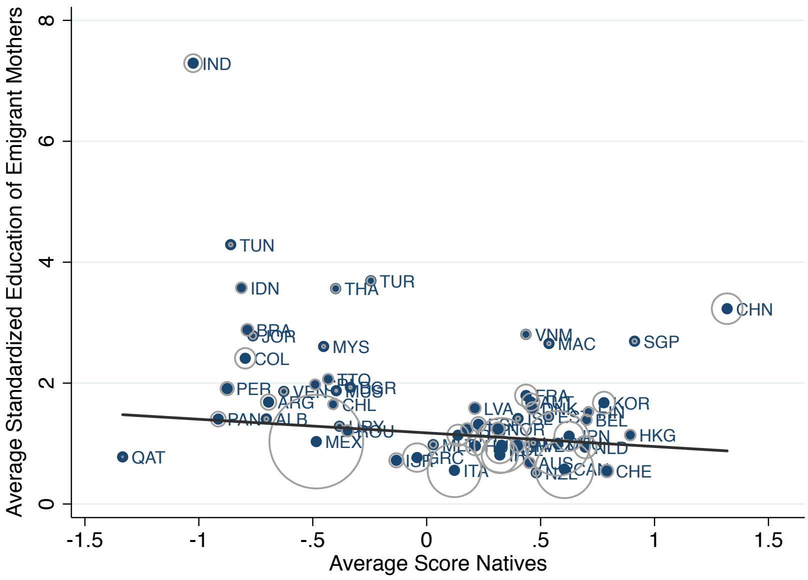
Overall, patterns of selection are on average remarkably similar between the PISA and the Census data. China, however, represents an interesting exception: while Chinese parents appear to be the most negatively selected in the PISA data, they turn out to be above the average country in terms of positive selection in the Census. This latter result is consistent with Feliciano (2005a), which shows that Chinese immigrants in the US are among the most positively selected in terms of education across countries of origin. The discrepancy between the two datasets is easily explained by the fact that the native population from China in PISA only includes people from Shanghai, which are substantially more educated compared to the rest of China. Therefore, while Chinese emigrants are positively selected compared to Chinese stayers as a whole, they are still negatively selected compared to stayers observed in the PISA sample.

Figure A.3: Possible Determinants of Emigrants' Selection



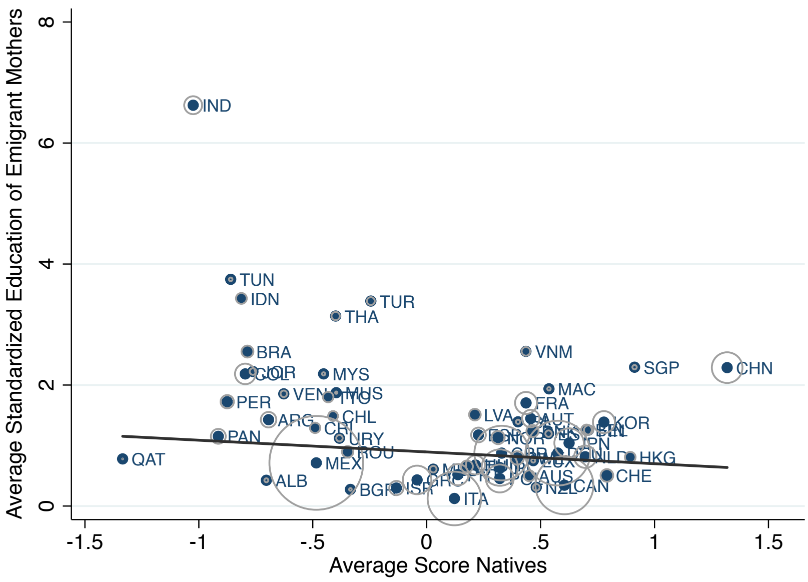
Notes: Each Panel plots the relationship between the average score among natives and a possible determinat of emigrants' selection. Panel (a) plots the difference between the average Gini Index faced by emigrants from country m in their respective host countries and the Gini Index in country m. Similarly, Panels (b) and (c) plot the difference between the average value faced by emigrants from country m and country m's value for the estimated return to education and the absolute income gap between the 80^{th} and the 20^{th} percentiles (in thousands of 2000 US dollars). Panel (d) plots the log real GDP per capita in 1985. Panel (e) and (f) plot the average across emigrants from m of the share of host country population born in country m and of the geographic distance between the host country and country m (in kilometers). The lines show the best linear fits.

Figure A.4: Selection on Parental Education



Notes: The Figure plots the average years of schooling of emigrant mothers from country m standardized by the average and the standard deviation of years of schooling of non-emigrant mothers in country m (y-axis) against the average PISA score of native students in country m (x-axis). The line shows the best (weighted) linear fit.

Figure A.5: Selection on Parental Education (Mothers Entirely Educated in Home Country)



Notes: The Figure plots the average years of schooling of emigrant mothers from country m standardized by the average and the standard deviation of years of schooling of non-emigrant mothers in country m (y-axis) against the average PISA score of native students in country m (x-axis). The sample includes only mothers entirely educated in their home country. The line shows the best (weighted) linear fit.

A.4 Additional Evidence on Mechanisms

A.4.1 Alternative Measures of Immigrants' Assimilation

In this section we show interaction results using different proxies for parents' integration in their host country. While we focus on years since migration in the main text, immigrants' assimilation is a complex process involving cultural, economic, linguistic and relational transitions, some of which might not be well captured by our proxy. At the same time, some dimensions of parents' assimilation considered here might be direct outcomes of their propensity to human capital accumulation, making the results harder to interpret.

We consider two sets of alternative measures. In Table A.18 we look at intermarriage with natives, which has been widely used as proxy for immigrants' assimilation (Gordon, 1964; Pagnini and Morgan, 1990), and is usually associated with favourable economic outcomes (Furtado and Trejo, 2013). In column 1 we interact a dummy identifying native fathers with T^m, and find that indeed the mother's country-of-origin effect is weaker when the father is a native. Column 2 shows that this patter is robust to the introduction of the other interaction terms explored in the main text. A possible complication arises from the fact that in these specifications we are not considering the father's country of origin, and if mothers from high PISA countries not matched to natives are systematically paired with fathers from high PISA country (and, chiefly, from their own same country), then the omission of a proxy for fathers' influence might explain the negative interaction. To explore this, in columns 3 and 4 we add to the previous specifications the average score from the father's country of origin. While the magnitude of our interaction of interest is unaffected, the coefficients are no longer statistically significant.

Table A.18: Heterogeneity with respect to Intermarriage with Natives

	Dependent Variable: No Grade Repeated			
	[1]	[2]	[3]	[4]
Score Country *m*	0.043***	0.138***	0.043***	0.149***
	(0.008)	(0.029)	(0.015)	(0.026)
Native Father × Score Country *m*	-0.031***	-0.019***	-0.032	-0.023
	(0.008)	(0.006)	(0.020)	(0.018)
Female	0.067***	0.067***	0.067***	0.067***
	(0.003)	(0.003)	(0.003)	(0.003)
Yrs Schooling Father	0.005***	0.005***	0.005***	0.005***
	(0.001)	(0.001)	(0.001)	(0.001)
Yrs Schooling Mother	0.006***	0.006***	0.007***	0.007***
	(0.001)	(0.001)	(0.001)	(0.001)
Yrs Since Migr Mother	0.001*	0.002***	0.001*	0.002***
	(0.001)	(0.000)	(0.001)	(0.000)
Log Family Income	0.032***	0.032***	0.029***	0.029***
	(0.008)	(0.008)	(0.008)	(0.008)
Score Country *m* × Yrs Since Migr Mother		-0.002***		-0.002***
		(0.001)		(0.001)
Score Country *m* × Yrs Schooling Mother		-0.006***		-0.006***
		(0.002)		(0.002)
Score Country *f*			-0.000	-0.006
			(0.018)	(0.018)
N	53081	53081	51428	51428
# Country *m*	61	61	61	61
R Squared	0.10	0.11	0.10	0.11
Comm Zone FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes

Notes: The Table shows results for second generation immigrants on the mother's side. *Score Country m* is the average math PISA score of natives (standardized to have mean 0 and standard deviation 1 across all countries participating to the test) in the country of birth of the mother, across all available waves. All specifications control for intercept, child age dummies, parents' age, family size, log family income, year fixed effect, (year-specific) quarter of birth fixed effect and father's immigrant status. Observations are weighted according to the provided sample weights. Standard errors are clustered by mother's country of origin. * denotes significance at 10%, ** at 5%, *** at 1%.

In Table A.19 we consider instead the measures of linguistic and cultural distances already discussed in the main text. The idea here is that parents which are linguistically or culturally far from the US norms are less likely to integrate, and perhaps to adapt to the locally prevalent practices and values in terms of children's education. Column 1 shows that indeed the gap between second generation immigrants from high and low PISA countries is larger when parents are linguistically far from the US, and column 2 confirms that this differential effect is robust to the inclusion of our baseline interactions. Results on cultural distance, while of the expected sign, are not statistically different from 0 (columns 3 and 4).

Table A.19: Heterogeneity with respect to Linguistic and Cultural Distance

	Dependent Variable: No Grade Repeated			
	[1]	[2]	[3]	[4]
Score Country *m*	0.033***	0.137***	0.032***	0.143***
	(0.006)	(0.028)	(0.007)	(0.027)
Score Country *m* × Mother Linguistic Distance	0.020***	0.017***		
	(0.006)	(0.006)		
Mother Linguistic Distance	-0.003	-0.002		
	(0.003)	(0.003)		
Father Linguistic Distance	-0.004*	-0.003		
	(0.002)	(0.002)		
Score Country *m* × Mother Cultural Distance			0.002	0.006
			(0.013)	(0.012)
Mother Cultural Distance			0.013	0.009
			(0.008)	(0.007)
Father Cultural Distance			-0.012***	-0.011***
			(0.003)	(0.003)
Female	0.067***	0.067***	0.067***	0.067***
	(0.003)	(0.003)	(0.003)	(0.003)
Yrs Schooling Father	0.005***	0.005***	0.005***	0.005***
	(0.001)	(0.001)	(0.001)	(0.001)
Yrs Schooling Mother	0.007***	0.007***	0.006***	0.007***
	(0.001)	(0.001)	(0.001)	(0.001)
Yrs Since Migr Mother	0.001	0.001***	0.001	0.001***
	(0.001)	(0.000)	(0.001)	(0.000)
Log Family Income	0.030***	0.029***	0.030***	0.030***
	(0.008)	(0.008)	(0.008)	(0.008)
Score Country *m* × Yrs Since Migr Mother		-0.002***		-0.002***
		(0.001)		(0.001)
Score Country *m* × Yrs Schooling Mother		-0.006***		-0.006***
		(0.002)		(0.002)
N	51420	51420	49438	49438
# Country *m*	60	60	47	47
R Squared	0.10	0.11	0.10	0.11
Comm Zone FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes

Notes: The Table shows results for second generation immigrants on the mother's side. *Score Country m* is the average math PISA score of natives (standardized to have mean 0 and standard deviation 1 across all countries participating to the test) in the country of birth of the mother, across all available waves. *Linguistic Distance* and *Cultural Distance* are standardized to take mean 0 and standard deviation 1 across all country pairs in the PISA sample (sources are discussed in the paper). All specifications control for intercept, child age dummies, parents' age, family size, log family income, year fixed effect, (year-specific) quarter of birth fixed effect and father's immigrant status. Observations are weighted according to the provided sample weights. Standard errors are clustered by mother's country of origin. * denotes significance at 10%, ** at 5%, *** at 1%.

Overall, the results in this section reinforce the message that the more immigrant parents integrate in the US the more the school performance of their children becomes similar across different countries of origin.

A.4.2 Immigrants' Ethnic Network

Throughout the paper, we stress the role of parents in the transmission of human capital, and we focus on second generation immigrants in order to fix the characteristics of the local environment. A potential complication arises from the fact that immigrant parents from different nationalities are likely to be differentially exposed to their own ethnic network, even within the same host country or region. The transmission of country-specific skills or attitudes towards education could also take place thorugh this channel, and the objective of this section is to investigate this possibility.

A.4.2.1 Borjas' Ethnic Capital

In his seminal work, Borjas (1992) uses data from the General Social Survey (GSS) and the National Longitudinal Survey of Youth (NLSY) to argue that the average level of education in the ethnic environment of parents, what he calls "ethnic capital", plays a role in the human capital accumulation process of the following generations in the US. To the extent that second generation immigrants from high-scoring countries are exposed to higher ethnic capital, this could represent a factor behind their superior performance at school additional to any direct interaction with their parents.

We use the Census data to construct a measure of the average years of education of parents of school-age children for each commuting zone and country of origin.[11] In Table A.20, we add this measure of ethnic capital as a control to our baseline specifications, shown in columns 1 and 3. No matter whether commuting zone fixed effects are introduced (column 4) or not (column 2), the coefficient on ethnic capital is positive and significant, consistently with Borjas' result. The coefficient on the average score of natives in the mother's country of origin is somewhat smaller in magnitude, but still positive and significant.

[11] We consider children between 8 and 15 years of age, consistently with the criterion used for our baseline sample. The results are similar when we use the same measure constructed at the state or the country level.

Table A.20: Ethnic Capital

	Dependent Variable: No Grade Repeated			
	[1]	[2]	[3]	[4]
Score Country *m*	0.045***	0.030***	0.028***	0.022**
	(0.013)	(0.008)	(0.009)	(0.009)
Ethnic Capital		0.008***		0.003**
		(0.002)		(0.002)
Female	0.068***	0.068***	0.068***	0.067***
	(0.003)	(0.003)	(0.003)	(0.003)
Mother Sec Edu	0.053***	0.041***	0.047***	0.043***
	(0.012)	(0.012)	(0.011)	(0.013)
Mother Ter Edu	0.057***	0.037***	0.054***	0.046***
	(0.012)	(0.013)	(0.010)	(0.012)
Father Sec Edu	0.042***	0.036***	0.036***	0.034***
	(0.013)	(0.012)	(0.010)	(0.010)
Father Ter Edu	0.061***	0.051***	0.058***	0.054***
	(0.015)	(0.012)	(0.011)	(0.011)
Log Family Income	0.042***	0.039***	0.036***	0.035***
	(0.010)	(0.009)	(0.008)	(0.008)
N	53081	53081	53081	53081
# Country *m*	61	61	61	61
R Squared	0.09	0.09	0.10	0.10
Comm Zone FE	No	No	Yes	Yes
Years Since Migr Mother	Yes	Yes	Yes	Yes

Notes: The Table shows results for second generation immigrants on the mother's side. *Score Country m* is the average math PISA score of natives (standardized to have mean 0 and standard deviation 1 across all countries participating to the test) in the country of birth of the mother, across all available waves. *Ethnic Capital* is the average years of education of all parents of 8- to 15-year-old children in the same commuting zone and born in country m. All specifications control for intercept, child age dummies, parents' age, family size, year fixed effect, (year-specific) quarter of birth fixed effect and father's immigrant status. Observations are weighted according to the provided sample weights. Standard errors are clustered by mother's country of origin. * denotes significance at 10%, ** at 5%, *** at 1%.

A.4.2.2 Horizontal and Obliqual Transmission

Parents' ethnic network could also play a role in the transmission of country-specific skills and attitudes towards education. Indeed, the literature on cultural transmission stresses the distinction between the vertical (parents to children), horizontal (children to children) and obliqual (other children's parents to children) transmission of cultural traits, all of which have been shown to be active in different settings (Bisin and Verdier, 2010). In the paper we stress the vertical channel, but to what extent does the performance of second generation immigrants also reflect the horizontal or obliqual ones?

To shed some light on this issue, we exploit the variation across commuting zones in the level of segregation across countries of origin. As discussed by Fernandez and Fogli (2009), local communities with a large share of individuals with the same ancestry might offer more opportunities for the horizontal or obliqual transmission of values through direct interaction, role models and punishments for behaviours not consistent with the social norm. If these channels are important, we would expect a larger country of origin-effect for parents located in such communities, as opposed to more isolated parents.

In Table A.21 we augment the baseline specification (shown in column 1) with interaction terms between T^m and measures of commuting zone-level segregation by country of origin. In particular, we consider the share of all (column 2), 35- to 45-year-old (column 3) and 8- to 15-year-old (column 4) residents born in country m. The coefficients for the interaction terms are positive for all specifications but marginally significant only for the second measure. Moreover, from the coefficient on T^m we can see that in all cases virtually the whole effect persists when the size of the ethnic network approaches zero. The gap in performance is therefore strong even when we focus on rather isolated parents, suggesting that our focus on the vertical channel of transmission might be well-warranted.

The results of Table A.21 should be interpreted with a caveat in mind. Several contributions to the cultural transmission literature argue that the prevalence of given cultural traits in the local context affects the incentives parents face when socializing their children, and that, depending on the setting, vertical and non-vertical (horizontal or obliqual) transmission might be either cultural substitutes or complements (Bisin and Verdier, 2010). Under cultural substitutability, it might be that parents that value education the most play a more active role in shaping human capital accumulation of their children when the horizontal transmission of positive attitudes towards education is muted, to same extent invalidating our interpretation of the results in Table A.21.

Table A.21: Heterogeneity with respect to the Segregation Rate

	Dependent Variable: No Grade Repeated			
	[1]	[2]	[3]	[4]
Score Country *m*	0.028***	0.025**	0.023**	0.027***
	(0.009)	(0.009)	(0.010)	(0.010)
Female	0.068***	0.068***	0.068***	0.068***
	(0.003)	(0.003)	(0.003)	(0.003)
Mother Sec Edu	0.047***	0.048***	0.048***	0.047***
	(0.011)	(0.012)	(0.012)	(0.012)
Mother Ter Edu	0.054***	0.054***	0.055***	0.054***
	(0.010)	(0.010)	(0.010)	(0.010)
Father Sec Edu	0.036***	0.036***	0.037***	0.036***
	(0.010)	(0.011)	(0.011)	(0.011)
Father Ter Edu	0.058***	0.059***	0.060***	0.058***
	(0.011)	(0.013)	(0.013)	(0.013)
Log Family Income	0.036***	0.036***	0.036***	0.036***
	(0.008)	(0.009)	(0.009)	(0.009)
Score Country *m* $\times$ Share from *m* (%)		0.006		
		(0.007)		
Share from *m*		0.003		
		(0.004)		
Score Country *m* $\times$ Share 35-45 from *m* (%)			0.009*	
			(0.005)	
Share 35-45 from *m*			0.006*	
			(0.003)	
Score Country *m* $\times$ Share 8-15 from *m* (%)				0.003
				(0.006)
Share 8-15 from *m*				0.002
				(0.005)
N	53081	53081	53081	53081
# Country *m*	61	61	61	61
R Squared	0.10	0.10	0.10	0.10
Comm Zone FE	Yes	Yes	Yes	Yes
Years Since Migr Mother	Yes	Yes	Yes	Yes

Notes: The Table shows results for second generation immigrants on the mother's side. *Score Country m* is the average math PISA score of natives (standardized to have mean 0 and standard deviation 1 across all countries participating to the test) in the country of birth of the mother, across all available waves. *Share from m*, *Share 35-45 from m* and *Share mothers from m* are, within the commuting zone of each student, the shares of, respectively, all residents, residents aged 35 to 45 and residents aged 8 to 15 born in country *m* (in percent). All specifications control for intercept, child age dummies, parents' age, number of siblings, log family income, year fixed effect, (year-specific) quarter of birth fixed effect and father's immigrant status. Observations are weighted according to the provided sample weights. Standard errors are clustered by mother's country of origin. * denotes significance at 10%, ** at 5%, *** at 1%.

In Table A.22 we explore this possibility by turning to the Time Use data, where we can observe a proxy for a parental $input$ for children's human capital accumulation. In particular, we ask whether the gap across nationalities in the time parents spend with their children is smaller when parents live in a community with a larger ethnic network from their country of origin, as it would be implied by cultural substitution. We construct a measure of segregation at the State and country of origin level, given by the share of residents in each State born in each country of origin (results with alternative measures of segregation along the lines of Table A.21 are very similar and available upon request). We then add to our baseline specification an interaction between this measure and T^p, the average score of natives in the country of origin of the interviewed parent. The results show that, if anything, the interaction term is positive, implying that parents spend more time with their children when living in a more segregated State. This finding is not consistent with a cultural substitution story, and provides further support for the fact that vertical transmission plays a key role.

Table A.22: Time Use - Heterogeneity with respect to the Segregation Rate

	Total	Educational	Recreational	Basic
	[1]	[2]	[3]	[4]
Score Country *p*	3.283	1.564	0.153	1.566
	(5.773)	(1.326)	(2.493)	(3.300)
Share from *p*	1.594	-0.007	1.702	-0.101
	(2.100)	(0.490)	(1.157)	(1.156)
Score Country *p* × Share from *p* (%)	4.510	0.359	4.073	0.078
	(4.621)	(0.959)	(2.488)	(2.430)
Parent Sec Edu	-3.009	4.285***	-3.132	-4.162**
	(5.111)	(0.540)	(3.079)	(2.014)
Parent Ter Edu	2.894	3.361***	-3.031	2.564
	(3.044)	(1.102)	(2.006)	(1.865)
Spouse Sec Edu	2.741	-1.911**	6.383**	-1.730
	(3.198)	(0.785)	(2.682)	(1.385)
Spouse Ter Edu	11.895***	2.079	6.890***	2.926
	(3.181)	(1.646)	(2.561)	(2.530)
Log Family Income	5.899***	0.597	-1.515	6.817***
	(2.102)	(0.658)	(0.977)	(1.314)
N	5659	5659	5659	5659
# Country *p*	59	59	59	59
Mean Dep. Var.	89.87	10.53	22.27	57.07
St. Dev. Dep. Var.	119.98	32.30	58.06	88.63
R Squared	0.24	0.06	0.10	0.22
State FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes

Notes: The sample includes only immigrant parents of children with at most 18 years. $Parent$ refers to the interviewed parent, $Spouse$ to the other one; $Mother$ is 1 when the interview parent is the mother. $Total$ refers to the total time spent in child care activities, while $Educational$, $Recreational$ and $Basic$ refer to the sub-categories defined in the text. *Score Country p* is the average math PISA score of natives (standardized to have mean 0 and standard deviation 1 across all countries participating to the test) in the country of birth of the interviewed parent, across all available waves. *Share from p* is the share of residents in the state where each parent lives born in country p (in percent). All specifications control for parents' age, number of children, number of male children, chidren's average age, years since migration, dummies for native spouses and for retired, full time students and disabled parents. Standard errors are clustered by the interviewed parent's country of origin. * denotes significance at 10%, ** at 5%, *** at 1%.

A.5 Complementarities

A.5.1 Complementarities between School and Parental Influence

Both our reduced form evidence and decomposition exercise are based on a specification where school inputs (as proxied by school fixed effects) and parental influence (as proxied by either the average score in parents' country of origin or country of origin fixed effects) are additively separable. However, a reasonable alternative would be having an interaction between these variables, capturing patterns of complementarity or substitutability between these two kinds of inputs. Were these interactions quantitatively important, the matching pattern between parents and schools of different "qualities" would become potentially important in explaining cross-country differences in the average PISA score, substantially complicating our counterfactual analysis.[12]

To assess the importance of this possibility, we allow for an interaction between the quality of school and parental inputs in our baseline reduced form specification. In particular, we use the average score among students with native parents in a given school as a proxy for school quality, and we ask whether the difference in performance between second generation immigrants from high and low PISA countries varies as a function of school quality.

Table A.23 shows our results. We find that the interaction term is small in magnitude and not significantly different from 0, no matter whether we use the school and country of origin PISA scores as baseline controls (column 2) or whether we absorb those in school and country of origin fixed effects. Moreover, the coefficient on T^m and the R^2 are virtually unaffected by the introduction of the interaction term (columns 2 and 4), suggesting that the linear specification is not missing much in terms of the fitting of the data.

[12]For example, in the case of complementarity between schooling and parental inputs, countries with a more assortative matching between parents and schools would obtain higher average scores.

Table A.23: Complementarities between School and Parental Influence - Reduced Form Results

	Dependent Variable: Math Test Score				
	[1]	[2]	[3]	[4]	[5]
Score Country *m*	0.327***	0.328***	0.243***	0.272***	
	(0.054)	(0.065)	(0.066)	(0.080)	
Score School *s*	0.777***	0.777***			
	(0.023)	(0.023)			
Score Country *m* × Score School *s*		-0.003		-0.041	0.016
		(0.038)		(0.042)	(0.035)
Female	-0.173***	-0.173***	-0.200***	-0.200***	-0.196***
	(0.019)	(0.019)	(0.022)	(0.022)	(0.023)
Mother Sec Edu	0.025	0.025	-0.013	-0.013	-0.029
	(0.035)	(0.035)	(0.039)	(0.039)	(0.034)
Mother Ter Edu	0.057*	0.057*	-0.006	-0.007	-0.030
	(0.036)	(0.036)	(0.042)	(0.042)	(0.040)
Father Sec Edu	0.000	0.000	0.009	0.009	0.007
	(0.031)	(0.031)	(0.022)	(0.022)	(0.022)
Father Ter Edu	0.005	0.005	0.041*	0.041*	0.037
	(0.039)	(0.038)	(0.031)	(0.031)	(0.031)
Mother Working × ISEI	0.004***	0.004***	0.003***	0.003***	0.003***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Father Working × ISEI	0.004***	0.004***	0.003***	0.003***	0.003***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
N	39573	39573	39573	39573	39573
# Country *m*	49	49	49	49	49
R Squared	0.51	0.51	0.67	0.67	0.67
Country *m* FE	No	No	No	No	Yes
School FE	No	No	Yes	Yes	Yes

Notes: The Table shows results for second generation immigrants on the mother's side. The sample includes only cases where both parents report a country of origin and the country of origin of the mother runs a PISA test on natives. *Score Country m* is the average math PISA score of natives (standardized to have mean 0 and standard deviation 1 across all countries participating to the test) in the country of birth of the mother, across all available waves. *Score School s* is the average math PISA score of students with both native parents in school s. All specifications control for intercept, students' age (in months), wave fixed effect and a dummy for father's immigrant status. Observations are weighted according to the provided sample weights. Standard errors are clustered by mother's country of origin, and inflated by the estimated measurement error in test scores. * denotes significance at 10%, ** at 5%, *** at 1%.

A.5.2 Complementarities between Maternal and Paternal Influence

Another possible form of complementarity is the one between the influence exterted by mothers and fathers. To explore this issue, in Table A.24 we focus on the sample of second generation immigrants with both parents born abroad, and augment the baseline specifications (reported in columns 1 and 3) with an interaction term between the average score in the mother's and father's country of origin. According to the results in columns 2 and 4, there seems some degree of positive complementarity between maternal and paternal influence, even though the coefficient of the interaction is not statistically significant.

Table A.24: Complementarities between Maternal and Paternal Influence - Reduced Form Results

	Dependent Variable: Math Test Score			
	[1]	[2]	[3]	[4]
Score Country *m*	0.273**	0.253**	0.181**	0.177*
	(0.107)	(0.116)	(0.089)	(0.091)
Score Country *f*	0.229*	0.187	0.017	0.005
	(0.123)	(0.140)	(0.090)	(0.096)
Score Country *m* × Score Country *f*		0.164		0.043
		(0.129)		(0.057)
Female	-0.142***	-0.140***	-0.205***	-0.205***
	(0.038)	(0.038)	(0.026)	(0.026)
Mother Sec Edu	0.079*	0.083*	-0.006	-0.006
	(0.044)	(0.045)	(0.044)	(0.044)
Mother Ter Edu	0.107*	0.112*	-0.032	-0.032
	(0.063)	(0.065)	(0.056)	(0.056)
Father Sec Edu	0.035	0.036	0.002	0.001
	(0.039)	(0.040)	(0.022)	(0.023)
Father Ter Edu	0.112**	0.107**	0.041	0.040
	(0.047)	(0.046)	(0.032)	(0.032)
Mother Working × ISEI	0.007***	0.007***	0.002**	0.002**
	(0.001)	(0.001)	(0.001)	(0.001)
Father Working × ISEI	0.006***	0.006***	0.002**	0.002**
	(0.001)	(0.001)	(0.001)	(0.001)
N	25534	25534	25534	25534
# Country *m*	47	47	47	47
# Country *f*	47	47	47	47
R Squared	0.31	0.31	0.70	0.70
Host Country FE	Yes	Yes	Yes	Yes
School FE	No	No	Yes	Yes

Notes: The Table shows results for second generation immigrants on both the mother's and the father's side. The sample includes only cases where both parents report a country of origin that runs a PISA test on natives. *Score Country m* and *Score Country f* are the average math PISA scores of natives (standardized to have mean 0 and standard deviation 1 across all countries participating to the test) in the country of birth of the mother and the father, across all available waves. All specifications control for intercept, students' age (in months) and wave fixed effect. Observations are weighted according to the provided sample weights. Standard errors are clustered by mother's country of origin, and inflated by the estimated measurement error in test scores. * denotes significance at 10%, ** at 5%, *** at 1%.

Ignoring this form of complementarity could lead to an understatement of the quantitative importance of parental influence for cross-country gaps in human capital. We investigate this possibility by replicating our decomposition exercise on the sub-sample of students which have parents of the same nationality. In this case, the country of origin fixed effect captures the combined effect of maternal and parental influence, including any complementarity between the two. Notably, this specification accomodates also the possibility that the degree of complementarity is country-of-origin-specific.

We estimate the regression

$$\begin{aligned} T^p_{icst} = \beta' ParentsEdu^p_{icst} + \lambda' ParentsOcc^p_{icst} + \gamma^p + \theta^p NatParents^p_{icst} + \\ + \rho' X^p_{icst} + \alpha_{cs} + \alpha_t + u^p_{icst} \end{aligned} \tag{A.1}$$

where the superscript p identifies the country of origin of both parents, γ^p is a parental country of origin fixed effect and $NatParents^p_{icst}$ is a dummy that takes value 1 for native parents. The parental component for country c is defined as

$$Parents^c = \gamma^p + \beta' \overline{ParentsEdu}^p + \lambda' \overline{ParentsOcc}^p \tag{A.2}$$

where $\overline{ParentsEdu}^p$ and $\overline{ParentsOcc}^p$ are within-country averages across native parents.

Table A.25 displays the resulting $V_{Parents}$ and V_{FE}, as defined in the main text. Consistently with the results in Table A.24, incorporating this form of complementarity amplifies the importance of parental influence. This is especially true for the specification with host country fixed effects, suggesting that the complementarity between mothers and fathers might be operative with particular relevance on the margin of school choice.

Table A.25: Decomposition with Complementarities between Maternal and Paternal Influence - Cross-Country Variance

	$V_{Parents}$ (%)		V_{FE} (%)	
	School FE	Host Country FE	School FE	Host Country FE
Unadjusted	24.24	51.61	23.84	49.87
Adjusted	20.96	48.36	20.58	46.72

Notes: The Table shows the ratio (in percent) between the cross-country variance of either the whole parental component ($V_{Parents}$) or the country specific intercept (V_{FE}) and the cross-country variance of the average math PISA score of natives. The sample only includes students with parents of the same nationality. Columns denoted by *School FE* (*Host Country FE*) refer to specifications that include school fixed effects (host country fixed effects). Adjusted variances are computed by subtracting the average squared standard errors (constructed using the provided replicate weights, and inflated by the estimated measurement error in test scores).

A.6 Development Accounting

In this section we investigate the implications of our results in terms of cross-country differences in output per worker. We start by decomposing the covariance between log GDP per worker and test scores between the contributions of the parental component and residual test scores; that is, we calculate

$$\frac{Cov(\log y_c, Parents^c)}{Cov(\log y_c, T^c)}$$

where y_c, is real GDP per worker, $Parents^c$ is the parental component as defined in the main text and T^c is the average performance of native students in country c. Table A.26 shows the results for both the school and host country fixed effects specifications. In both cases, around 12% of the covariance is driven by the parental component. The fact that this figure is smaller than the relative variance displayed in Table 1.5 in the main text reflects that the parental component, besides varying less across countries compared to the residual test score, is also slightly less correlated with other factors of production. In other words, parental influence is more equally distributed across rich and poor countries relative to other forms of physical and human capital.

Table A.26: Decomposition Results - Covariance with GDP per worker

	School FE	Host Country FE
$\frac{Cov(\log y_c, Parents^c)}{Cov(\log y_c, T^c)}$ (%)	11.60	12.79

Notes: The Table shows the ratio (in percent) between the cross-country covariance between log GDP per worker and the parental component and the cross-country covariance between log GDP per worker and average test scores of native students. Columns denoted by *School FE* (*Host Country FE*) refer to specifications that include school fixed effects (host country fixed effects).

To have a sense of the absolute magnitude of the contribution of each component, we implement a simple development accounting exercise. We follow Klenow and Rodríguez-Clare (1997) and much of the literature in postulating an aggregate Cobb-Douglas production function which can be written in per worker terms as:

$$y_c = A_c \left(\frac{k_c}{y_c}\right)^{\frac{\alpha}{1-\alpha}} h_c$$

This formulation allows an additive decomposition of the variance of $\log y_c$ into the contributions of the covariances between $\log y_c$ and the appropriately weighted covariances of the logs of TFP, capital to output ratio and human capital. We are interested in the magnitude and the composition of the latter term,

$$\frac{Cov(\log y_c, \log h_c)}{Var(\log y_c)}$$

which represents our measure of the overall contribution of human capital. We assume that human capital per worker is given by the exponential form

$$h_c = \exp\{\beta_s s_c + \beta_t T^c\}$$

where s_c is average years of schooling in country c. For our baseline, we follow Hanushek and Woessmann (2012b) in setting $\beta_s = 0.1$ and $\beta_t = 0.2$, which are picked to match estimates of the returns to schooling and test performance in the labour market.

Table A.27 shows that the baseline measure of human capital accounts for 29% of the variation in GDP per capita. Differences in years of schooling are responsible for 19% of the variance, while test scores account for the remaining 10%. Out of this 10%, a little bit more of a percentage point is to be attributed to the parental component.

Table A.27: Development Accounting - Results

	Baseline	$\beta_t = 0$	$\beta_s = 0$	$\beta_s = 0, T^c = Parents^c$	
				School FE	Host Country FE
$\frac{Cov(\log y_c, \log h_c)}{Var(\log y_c)}$ (%)	29.32	19.39	9.93	1.15	1.27

Notes: The Table shows the ratio (in percent) between the cross-country covariance between log GDP per worker and log human capital per worker and the variance of log GDP per worker. Each column corresponds to a different specification for h_c. Columns denoted by *School FE* (*Host Country FE*) refer to specifications that include school fixed effects (host country fixed effects).

Overall, the results in this section suggest that parental influence does not account for a sizable portion of the cross-country variation in GDP. This is due to the joint effect of two patterns: first, the variation in parental influence is not large enough and, second, parental influence does not covary strongly with other factors of production.

Appendix B

Appendix to Chapter 2

B.1 Derivations

B.1.1 Probability of attending college

The probability that an agent with family income y and ability z goes to college is given by

$$
\begin{aligned}
P[S(y,z)=1] &= P\left[\frac{\bar{\eta} r_H}{[1+\tau(y,z)]^{\eta}}\varepsilon_H > r_L z^{\alpha}\varepsilon_L\right] \\
&= P\left[\varepsilon_L < \frac{\bar{\eta} r_H}{[1+\tau(y,z)]^{\eta} r_L z^{\alpha}}\varepsilon_H\right] \\
&= \int_0^{\infty} F_H\left(\frac{\bar{\eta} r_H}{[1+\tau(y,z)]^{\eta} r_L z^{\alpha}}\varepsilon, \varepsilon\right) \mathrm{d}\varepsilon
\end{aligned}
$$

where $F_H(\varepsilon_L, \varepsilon_H)$ is the derivative of the joint distribution of ε_L and ε_H with respect to ε_H. Given that

$$F_H(\varepsilon_L, \varepsilon_H) = \theta \varepsilon_H^{-\theta-1} \exp\left\{-\varepsilon_L^{-\theta} - \varepsilon_H^{-\theta}\right\}$$

we have

$$F_H\left(\frac{\bar{\eta} r_H}{[1+\tau(y,z)]^{\eta} r_L z^{\alpha}}\varepsilon, \varepsilon\right) = \theta \varepsilon^{-\theta-1} \exp\left\{-\varepsilon^{-\theta}\left(\frac{(\bar{\eta} r_H)^{\theta} + \left([1+\tau(y,z)]^{\eta} r_L z^{\alpha}\right)^{\theta}}{(\bar{\eta} r_H)^{\theta}}\right)\right\}$$

Substituting in the integral above we get

$$
\begin{aligned}
P[S(y,z)=1] &= \int_0^\infty \theta\varepsilon^{-\theta-1}\text{exp}\left\{-\varepsilon^{-\theta}\left(\frac{(\bar{\eta}r_H)^\theta+\left([1+\tau(y,z)]^\eta\, r_L z^\alpha\right)^\theta}{(\bar{\eta}r_H)^\theta}\right)\right\}\mathrm{d}\varepsilon \\
&= \frac{(\bar{\eta}r_H)^\theta}{(\bar{\eta}r_H)^\theta+\left([1+\tau(y,z)]^\eta\, r_L z^\alpha\right)^\theta}\int_0^\infty \theta\left(\frac{(\bar{\eta}r_H)^\theta+\left([1+\tau(y,z)]^\eta\, r_L z^\alpha\right)^\theta}{(\bar{\eta}r_H)^\theta}\right)\varepsilon^{-\theta-1}* \\
&\quad * \exp\left\{-\varepsilon^{-\theta}\left(\frac{(\bar{\eta}r_H)^\theta+\left([1+\tau(y,z)]^\eta\, r_L z^\alpha\right)^\theta}{(\bar{\eta}r_H)^\theta}\right)\right\}\mathrm{d}\varepsilon
\end{aligned}
$$

Solving the integral,

$$
\begin{aligned}
P[S(y,z)=1] &= \frac{(\bar{\eta}r_H)^\theta}{(\bar{\eta}r_H)^\theta+\left([1+\tau(y,z)]^\eta\, r_L z^\alpha\right)^\theta}\left[\exp\left\{-\varepsilon^{-\theta}\left(\frac{(\bar{\eta}r_H)^\theta+\left([1+\tau(y,z)]^\eta\, r_L z^\alpha\right)^\theta}{(\bar{\eta}r_H)^\theta}\right)\right\}\right]_0^\infty \\
&= \frac{(\bar{\eta}r_H)^\theta}{(\bar{\eta}r_H)^\theta+\left([1+\tau(y,z)]^\eta\, r_L z^\alpha\right)^\theta}\left[1-0\right] \\
&= \frac{(\bar{\eta}r_H)^\theta}{(\bar{\eta}r_H)^\theta+\left([1+\tau(y,z)]^\eta\, r_L z^\alpha\right)^\theta}
\end{aligned}
$$

The probability of not attending college is just the complement of this,

$$
P[S(y,z)=0] = \frac{\left([1+\tau(y,z)]^\eta\, r_L z^\alpha\right)^\theta}{(\bar{\eta}r_H)^\theta+\left([1+\tau(y,z)]^\eta\, r_L z^\alpha\right)^\theta}
$$

B.1.2 Average wages

In order to compute average wages, I need to derive an expression for the expected value of the idiosyncratic shock for individuals that choose a given occupation. For the high skill sector,

$$
\begin{aligned}
P[\varepsilon_H < x | S(y,z)=1] &= \frac{P[\varepsilon_H < x \wedge \varepsilon_L < \frac{\bar{\eta}r_H}{[1+\tau(y,z)]^\eta r_L z^\alpha}\varepsilon_H]}{P[S(y,z)=1]} \\
&= \frac{1}{P[S(y,z)=1]}\int_0^x F\left(\frac{\bar{\eta}r_H}{[1+\tau(y,z)]^\eta\, r_L z^\alpha}\varepsilon\right) f(\varepsilon)\mathrm{d}\varepsilon \\
&= \exp\left\{-x^{-\theta}\left(\frac{(\bar{\eta}r_H)^\theta+\left([1+\tau(y,z)]^\eta\, r_L z^\alpha\right)^\theta}{(\bar{\eta}r_H)^\theta}\right)\right\} \\
&= G_H(x)
\end{aligned}
$$

That is, the idiosyncratic shock for agents choosing the high skill sector follows a Frechet distribution with shape parameter θ and scale parameter $\left(\frac{(\bar{\eta}r_H)^\theta+\left([1+\tau(y,z)]^\eta r_L z^\alpha\right)^\theta}{(\bar{\eta}r_H)^\theta}\right)^{\frac{1}{\theta}}$. The expected value of this random variable is

$$\mathbb{E}\left[\varepsilon_H | S(y,z)=1\right] = \left(\frac{(\bar{\eta}r_H)^\theta + \left([1+\tau(y,z)]^\eta r_L z^\alpha\right)^\theta}{(\bar{\eta}r_H)^\theta}\right)^{\frac{1}{\theta}} \Gamma\left(1-\frac{1}{\theta}\right) \tag{B.1}$$

Plugging (B.1) in (2.8), we obtain

$$\begin{aligned}
w_H(y,z) &= \frac{\bar{\eta}(1+\eta)r_H}{(1+\tau(y,z))^\eta} E[\epsilon_H(i)|S(y,z)=1] \\
&= (1+\eta)\left[\left(\frac{\bar{\eta}r_H}{(1+\tau(y,z))^\eta}\right)^\theta + (r_L z^\alpha)^\theta\right]^{\frac{1}{\theta}} \Gamma\left(1-\frac{1}{\theta}\right)
\end{aligned}$$

Similarly, for the low skill sector

$$\begin{aligned}
P[\varepsilon_L < x | S(y,z)=0] &= \frac{P[\varepsilon_L < x \wedge \varepsilon_H < \frac{[1+\tau(y,z)]^\eta r_L z^\alpha}{\bar{\eta}r_H}\varepsilon_L]}{P[S(y,z)=0]} \\
&= \frac{1}{P[S(y,z)=0]} \int_0^x F\left(\frac{[1+\tau(y,z)]^\eta r_L z^\alpha}{\bar{\eta}r_H}\varepsilon\right) f(\varepsilon)\mathrm{d}\varepsilon \\
&= \exp\left\{-x^{-\theta}\left(\frac{(\bar{\eta}r_H)^\theta + \left([1+\tau(y,z)]^\eta r_L z^\alpha\right)^\theta}{\left([1+\tau(y,z)]^\eta r_L z^\alpha\right)^\theta}\right)\right\} \\
&= G_L(x)
\end{aligned}$$

The idiosincratic shock for agents employed in the low skill sector follows a Frechet distribution with shape parameter θ and scale parameter $\left(\frac{(\bar{\eta}r_H)^\theta + ([1+\tau(y,z)]^\eta r_L z^\alpha)^\theta}{([1+\tau(y,z)]^\eta r_L z^\alpha)^\theta}\right)^{\frac{1}{\theta}}$. The expected value of this random variable is

$$\mathbb{E}\left[\varepsilon_L | S(y,z)=0\right] = \left(\frac{(\bar{\eta}r_H)^\theta + \left([1+\tau(y,z)]^\eta r_L z^\alpha\right)^\theta}{\left([1+\tau(y,z)]^\eta r_L z^\alpha\right)^\theta}\right)^{\frac{1}{\theta}} \Gamma\left(1-\frac{1}{\theta}\right) \tag{B.2}$$

Plugging (B.2) in (2.7), we obtain

$$\begin{aligned}
w_L(y,z) &= r_L z^\alpha E[\epsilon_L(i)|S(y,z)=0] \\
&= \left[\left(\frac{\bar{\eta}r_H}{(1+\tau(y,z))^\eta}\right)^\theta + (r_L z^\alpha)^\theta\right]^{\frac{1}{\theta}} \Gamma\left(1-\frac{1}{\theta}\right)
\end{aligned}$$

B.2 Competitive Equilibrium

A competitive equilibrium is given by a set of agents' choices on college attendance $\{S(i)\}_{i\in N}$, investment in educational units $\{e(i)\}_{i\in N}$, total efficiency units $\{L,H\}$ and factor prices $\{r_L, r_H\}$ such that

- Conditional on attending college, agents choose $e(i)$ to maximize their consumption net of college costs
- Agents make the college attendance choice that maximizes their utility, taking prices as given
- A representative firm hires low and high skill efficiency units to maximize profits, taking prices as given

$$\max_{H,L} A\left[L^{\rho} + BH^{\rho}\right]^{\frac{1}{\rho}} - r_L L - r_H H$$

- r_L and r_H clear the low skilled and high skilled labor market respectively,

$$L = \sum_i z(i)^{\alpha} \varepsilon_L(i)$$

$$H = \sum_i e(i)^{\eta} \varepsilon_H(i)$$

B.3 Algorithm for the Counterfactual in Section 6.3

To implement the counterfactual analysis when $\tau(y,z)$ is allowed to depend linearly on z, I follow these steps

1. Estimate r_L, r_H, α, η, α_1, β_1, α_2, β_2, α_3 and β_3 from (2.12) by NLS
2. Back out the realization of ε_H for those employed in the high skill sector and the realization of ε_L for those employed in the low skill sector from

$$\widehat{\varepsilon}_L(i) = \frac{w_L(i)}{\widehat{r}_L z^{\widehat{\alpha}}}$$

$$\widehat{\varepsilon}_H(i) = \frac{w_H(i)}{\widehat{r}_H \widehat{\bar{\eta}}(1+\widehat{\eta})(1+\widehat{\tau}(y,z))^{\widehat{\eta}}}$$

where $\widehat{r}_L$, $\widehat{r}_H$, $\widehat{\alpha}$, $\widehat{\eta}$ and $\widehat{\tau}(y,z)$ are the estimates coming from Step 1.

3. Compute H and L by summing up all the efficiency units (weighted by the provided sample weights) supplied by all individuals in the two sectors
4. Back out B using

$$B = \frac{\widehat{r}_H}{\widehat{r}_L}\left(\frac{H}{L}\right)^{1-\rho}$$

5. For each individual attending college, draw a $\varepsilon_L(i)$ shock from the truncated Frechet distribution with CDF

$$\begin{aligned} P[\varepsilon_L(i) < x | S(i) = 1] &= P[\varepsilon_L(i) < x | \varepsilon_L(i) < \frac{\bar{\eta} r_H}{\left[1+\tau(y,z)\right]^{\eta} r_L z^{\alpha}} \widehat{\varepsilon}_H(i)] \\ &= \exp\left\{-x^{-\theta} + \left(\frac{\bar{\eta} r_H \widehat{\varepsilon}_H(i)}{\left[1+\tau(y,z)\right]^{\eta} r_L z^{\alpha}}\right)^{-\theta}\right\} \end{aligned}$$

Similarly, for each individual not attending college, draw a $\varepsilon_H(i)$ shock from the truncated Frechet distribution with CDF

$$
\begin{aligned}
P[\varepsilon_H(i) < x | S(i) = 0] &= P[\varepsilon_H(i) < x | \varepsilon_H(i) < \frac{[1+\tau(y,z)]^{\eta}\, r_L z^{\alpha}}{\bar{\eta} r_H} \widehat{\varepsilon}_L(i)] \\
&= \exp\left\{ -x^{-\theta} + \left(\frac{[1+\tau(y,z)]^{\eta}\, r_L z^{\alpha}}{\bar{\eta} r_H} \right)^{-\theta} \right\}
\end{aligned}
$$

6. Determine the counterfactual college attendance choices for all individuals setting $\tau^C(y,z) = \widehat{\alpha}_3 + \widehat{\beta}_3 z$ and the corresponding changes in output and wages. In order to compute the new counterfactual equilibrium, first guess a value for $\frac{H}{L}$ and the update until convergence

7. Repeat Steps 5-6 1000 times and compute the 5^{th} and 95^{th} percentiles of the distribution of each counterfactual outcome