Immigration and Occupational Comparative Advantage*

Gordon Hanson

Harvard and NBER

Chen Liu

National University of Singapore

June 8, 2021

Abstract

Job choice by high-skilled foreign-born workers in the US correlates strongly with country of origin. We apply a Fréchet-Roy model of occupational choice to evaluate the causes of immigrant sorting. In a gravity specification, we find that revealed comparative advantage in the US is stronger for workers from countries with higher education quality in occupations that are more intensive in cognitive reasoning, and for workers from countries that are more linguistically similar to the US in occupations that are more intensive in communication. Our findings hold for immigrants who arrived in the US at age 18 or older (who received their K-12 education abroad) but not for immigrants who arrived in the US as children (who received their K-12 education domestically). We obtain similar results for immigrant sorting in Canada, which supports our interpretation that origin-country education quality, rather than US immigration policy, is what drives sorting patterns. In counterfactual analysis, we evaluate the consequences of reallocating visas for college-educated immigrants according to origin-country education quality.

*Contact information: Gordon Hanson, gordon_hanson@hks.harvard.edu; Chen Liu: ecsliuc@nus.edu.sg.

1 Introduction

The sorting of foreign-born workers across occupations in the US labor market correlates strongly with country of origin (Patel and Vella, 2013; Hanson and Liu, 2017). Consider immigrants from China and the Philippines, who in 2016 respectively accounted for a similar 1.5% and 1.2% of the prime-age US population with at least a college education. Yet, they represented 12.4% versus 1.2% of those employed as medical scientists and 0.9% versus 11.8% of those employed as licensed practical nurses.[1] Such differential sorting is common (see Appendix Table A.1). Among college-educated US residents, there is strong over-representation of workers from India in computer programming, from South Korea in dental health, and from Pakistan in service-station management, among many similar examples.[2]

There is a long history of immigrant groups concentrating in particular lines of work. Recent literature, for instance, documents Vietnamese immigrants specializing as manicurists (Federman et al., 2006), Mexican immigrants working as manual laborers and farm workers (Munshi, 2003; Woodruff and Zenteno, 2007), and Haitian immigrants serving as taxi drivers (Jackson and Schneider, 2011). Such clustering is consistent with the presence of migration networks, which connect less-educated immigrants to specific jobs and thereby reduce their costs of entry into a new and unfamiliar labor market (McManus, 1990; Munshi, 2020). What is distinct about the examples cited above is that the jobs involved—medical research, licensed nursing, business management, computer programming—typically require a college degree.[3] Having attained an advanced education and selected into immigration, individuals then appear to choose occupations based in part on where they were born.[4]

In this paper, we use a Fréchet-Roy (1951) framework to evaluate the causes of immigrant sorting across jobs. Perhaps the most obvious force behind sorting is language. When working in the US, immigrants from non-English-speaking countries may have a comparative disadvantage in jobs that are intensive in communication (Peri and Sparber, 2009; Oreopoulos, 2011). A second force, also related to comparative advantage, is the quality of education in the origin country. Countries that excel in math and science training, for instance, may be more likely to produce immigrants who work in STEM fields (Hunt and Gauthier-Loiselle, 2010). A third force behind sorting is the cost of migration. If policy barriers in a destination

[1]These figures are from the 2014-2018 American Communities Survey. Prime-age workers are those 25 to 54 years old; having a college education means having completed at least four years of college.

[2]Over 2014-2018, immigrants from India were 3.2% of the US college-educated population 25 to 54 years old and 15.5% of computer programmers, immigrants from Korea were 0.9% of this population and 6.0% of dental technicians, and immigrants from Pakistan were 0.3% of this population and 5.7% of service-station managers.

[3]See Cortés and Pan (2014) on the labor-market consequences of inflows of female nurses from the Philippines for US native-born workers. In our analysis, we study male workers only.

[4]In the literature, the sorting of less-educated immigrants into specific occupations has a strong spatial dimension, which suggests that localized immigrant job networks, as opposed to origin-country-specific training or labor-market experience, are behind national clustering patterns (Patel and Vella, 2013). By contrast, the instances of sorting among the college-educated foreign born that we document are a US-wide phenomenon.

differ across origin countries, then immigrants from high-migration-cost countries may be more positively selected in terms of skill (Grogger and Hanson, 2011).[5]

The first step in our analysis is to apply a Fréchet-Roy model (e.g., Burstein et al., 2015; Hsieh et al., 2019) to derive a gravity expression for the share of college-educated immigrants from a given origin country who are employed in a given US occupation. These employment shares are a function of wages at the occupation level in the US, and labor productivity, migration costs, and alternative employment opportunities for workers from the origin. By pooling data across origin-country groups and occupations in the US and regressing log employment shares on country-of-origin and occupation fixed effects, we neutralize the importance of alternative employment options, for each national-origin group, and average US wages, for each occupation, in occupational choice.[6] The residual component of these employment shares, which is our focus, is revealed occupational comparative advantage.

In theory, origin-country revealed comparative advantage by occupation is determined by fundamental comparative advantage and migration costs.[7] For more-educated workers, fundamental comparative advantage is likely shaped by the quality of a country's educational institutions, where the skills that these institutions impart may vary in importance across jobs (Hunt, 2011; Peri et al., 2015). Russia, whose universities excel in specific subfields of mathematics, produces mathematicians who in turn excel in these fields when competing against their US counterparts (Borjas and Doran, 2012). However, its workers may not be equipped to succeed in US managerial or sales-related professions. In a similar vein, whereas on average foreign-born workers with stronger English language skills achieve higher earnings in the US labor market (Dustmann and Fabbri, 2003; Bleakley and Chin, 2004; Hunt, 2015), the return to language ability may be larger in jobs that require stronger communication skills (e.g., management, sales) than in those that do not (e.g., engineering, mathematics) (Chiswick and Taengnoi, 2007). Migration costs may also affect occupational sorting by favoring migrants with higher earnings potential (Lazear, 2021).

Motivated by analysis of comparative advantage in international trade, we model occupational comparative advantage by interacting origin-country attributes with occupation-specific job requirements.[8] Origin country attributes include linguistic similarity with the US (Isphording and Otten, 2014; Melitz and Toubal, 2014), national performance in standardized exams conducted by the Program for International Student Assessment (PISA)

[5]Related work addresses native-born educational and occupational responses to immigration (Llull, 2018; Burstein et al., 2020), and the tendency of foreign-born workers to experience occupational downgrading in the destination country (Dustmann et al., 2013).

[6]This double-differencing is akin to that used in estimating the gravity model of trade. See, e.g., Head and Ries (2001) and Novy (2013) on the use of double differencing to identify the components of trade costs.

[7]In the Fréchet-Roy framework, fundamental comparative advantage is defined by the location parameters of the Fréchet distributions for worker productivity in the origin countries under consideration.

[8]See, e.g., Romalis (2004), Levchenko (2007), and Chor (2010) for alternative approaches to interacting country characteristics (e.g., factor endowments, quality of institutions) and industry characteristics (e.g., factor intensity, product complexity) to account for comparative advantage when estimating gravity models of trade.

(Hanushek and Kimko, 2000; Hanushek and Woessmann, 2011), and geographic distance to the US (Head and Mayer, 2013). Occupational requirements including performing tasks that require cognitive reasoning, interpersonal communication, manual effort, and repetitive operations (Autor et al., 2003; Deming, 2017). We control for bilateral migration costs not absorbed by origin-country fixed effects via additional interactions.

Our specification will capture the role of comparative advantage in immigrant occupational sorting if workers from countries that score relatively highly in math and science exams excel in jobs that require stronger cognitive skills and if workers from English-speaking countries perform relatively well in jobs that require greater face-to-face interaction with customers or co-workers. A substantial body of literature uses PISA exam scores to evaluate the contribution of education quality and cognitive skills to economic development (see, e.g., Hanushek and Woessmann, 2011; Woessmann, 2016).[9] Our contribution is to show how origin-country education quality affects the matching of workers to jobs abroad and thereby helps determine the composition of international labor flows.

A vibrant body of work uses the task intensity of jobs to evaluate how specialization in routine operations affects worker exposure to technological change (e.g., Autor and Dorn, 2013; Goos et al., 2014; Hershbein and Kahn, 2018), how the labor-market return to cognitive and social skills has evolved over time (e.g., Beaudry et al., 2016; Deming, 2017; Deming and Noray, 2020), and how immigration affects US labor-market outcomes (Peri and Sparber, 2009, 2011), among other topics. We show how task intensity affects the matching of more-educated foreign-born workers to occupations and thus determines the intensity of competition at the high end of the labor market.

More broadly, our work connects to literature on worker sorting. In theory, the matching of more-skilled workers to more skill-intensive tasks determines worker exposure to international competition (Costinot and Vogel, 2010) and technological change, where skill may be uni-dimensional (Acemoglu and Autor, 2011) or multi-dimensional (Lindenlaub, 2017). In a Fréchet-Roy context, positive sorting explains cross-country differences in agricultural productivity (Lagakos and Waugh, 2013), the consequences of reduced gender and racial discrimination (Hsieh et al., 2019), and which native-born workers are most affected by immigration (Burstein et al., 2020). We identify the country characteristics (cognitive, linguistic skills) that complement occupational requirements (cognitive reasoning, interpersonal communication) in the sorting process, and test for positive sorting in the context of multi-dimensional skills. The empirical mapping that we uncover pins down how changing the skill bias of immigration policy would alter occupational employment and wages.

[9] Also on the use of PISA exam scores, see Lavy (2015) and Hanushek et al. (2020) on how cross-country differences in student preferences and behavior affect test outcomes; Akyol et al. (2018) and Mogstad et al. (2020) on constructing country performance rankings when using noisy measures of individual ability; and Xiang and Yeaple (2018) on estimating the magnitudes of cognitive and non-cognitive labor skills by country.

Consistent with Fréchet-Roy logic, we find that revealed comparative advantage (defined relative to the average occupation) is stronger in occupations more intensive in cognitive reasoning for countries with higher PISA test scores and in occupations more intensive in interpersonal communication for countries that are more linguistically similar to the US. Comparing countries at the 25^{th} and 75^{th} percentiles of PISA math scores, the higher scoring country would have a 87% higher share of its college-educated immigrant labor in US financial management jobs, which is the occupation at the 75^{th} percentile of intensity in cognitive tasks. Similarly, when comparing countries at the 25^{th} and 75^{th} percentiles of linguistic similarity with the US, the higher scoring country would have a 21% higher share of its college-educated immigrant labor in US executive management positions, which is the occupation at the 75^{th} percentile of intensity in communication tasks.[10]

In interpreting our results as capturing how comparative advantage affects immigrant sorting across jobs, we confront several potential confounds. One is origin-country bias in US immigration policy. If visas for more-educated workers were awarded on the basis of skill alone, then immigrant sorting across jobs would largely reflect origin-country capabilities: immigrants from countries with higher quality education would concentrate in jobs requiring more analytical reasoning, and those from countries with greater facility in English would concentrate in jobs more reliant on communication skills. In reality, since 2000 the US has granted 65% of permanent residence visas (green cards) to immigrants who are sponsored by family members in the US and just 14% to immigrants sponsored by US employers, with a further 13% going to refugees and asylees and 7% awarded based on other criteria (OIS, 2020). Although college-educated immigrants are more likely to enter the US on employment visas than are the non-college educated (Bound et al., 2017; Gelatt, 2020), the intricate system of preferences used to award green cards means that we do not know whether or not visa allocation rules somehow distort comparative advantage.[11]

To evaluate the potentially confounding effects of US visa policies, we re-estimate our specifications using data on the allocation of foreign and native-born labor across occupations in Canada. Distinct from the US, Canada allocates immigration visas using a point system, which favors working-age individuals with more education, stronger job skills, and facility in English or French (Antecol et al., 2003; Schoellman, 2012). Our results for Canada are very similar to our results for the US, which suggests that our findings are not a byproduct of US immigration policy and instead capture origin-country occupational strengths. Despite its complexity, US immigrant policy does not suppress the role of comparative advantage in how foreign-born workers sort across jobs.

[10]For evidence of how the composition of migrants affects regional comparative advantage, see Pellegrina and Sotelo (2021) and Hanson (2021).

[11]Data from the New Immigrant Survey for 2003 show that the fraction of immigrants with a college education or advanced degree is 82% of those entering on employer-sponsored visas, 37% of those entering as spouses of US residents, and 33% of those entering as refugees or asylees (Gelatt, 2020).

A second potential confound is the impact of culture and social norms on migrant behavior (e.g., Fernandez and Fogli, 2009). If families of certain ancestries place higher value, say, on having a job in a STEM field, then higher standardized test scores could reflect these values, rather than origin-country education quality. To isolate the impact of being educated in the origin country, we separate the foreign born into those who arrived in the US at age 18 or older, and therefore likely completed their K-12 education in their country of birth, and those who arrived before age 18 (or before age 13), and therefore were educated partly or entirely in the US. Our results on comparative advantage are much weaker for the under-18 (and under-13) arrivals than for the 18-and-older arrivals. When we further limit the sample to US-born individuals and separate them according to their ancestral country of origin, we again find a weak relationship between origin-country comparative advantage and occupational sorting by ancestry. These results suggest that exposure to origin-country educational institutions shapes the occupational sorting of migrant labor abroad.

As a final exercise, we evaluate the general equilibrium implications of alternative policies for allocating visas under immigrant sorting by comparative advantage. We perform counterfactual exercises in which we artificially reallocate visas from immigrants with low PISA scores to immigrants with high PISA scores. These rules in turn alter the composition of immigrants by country of origin and thus the relative supply of labor across US occupations. Were the US to favor individuals with high PISA scores in the allocation of visas—in the spirit of Canada's point system for immigration—immigration from East Asia would increase while immigration from Latin America would decrease. The consequence would be an expansion in US employment in STEM-related occupations and a contraction in employment in most other occupations. Wages at the occupational level would fall in line with expansions in employment. Implications for native-born wages depend on whether foreign and native-born workers are perfect or imperfect substitutes in employment.

In Section 2, we outline our theoretical and empirical framework; in Section 3, we describe our data and patterns of immigrant sorting by occupation; in Sections 4 and 5, we present our empirical findings and counterfactual analysis; and in Section 6, we conclude.

2 Model

We use a Roy model of occupation choice to derive our estimating equation and for counterfactual analysis. In the model, workers from each origin country s choose a destination country d in which to reside and an occupation o in which to work. The model structure implies that the allocation of workers across destinations and occupations is a function of national comparative advantage and bilateral migration costs. Aggregating across these allocations for a given destination country (e.g., the US) allows us to evaluate counterfactually how changing bilateral migration costs would affect earnings in the destination.

2.1 Labor supply

Consider the migration and occupation choices for labor-market group G_s (e.g., college-educated men born in China). Within each group, there is a continuum of workers i. Each worker independently draws a vector of productivity values $a^{i,s}_{d,o}$ across destination-countries d and occupations o from a univariate Fréchet distribution, which has the following CDF,

$$F\Big(a^{i,s}_{d,o} \leqslant x\Big) = \exp\Bigg\{-\Big(\frac{x}{T^s_{d,o}}\Big)^{-\theta}\Bigg\}, \quad \theta > 1.$$ [12]

The distribution has two parameters, θ and $T^s_{d,o}$. With $T^s_{d,o}$ fixed, a larger θ corresponds to a smaller within-group dispersion of productivity. $T^s_{d,o}$ is the scale parameter, where a larger value corresponds to a higher level of average group productivity and greater within-group productivity dispersion. With the dispersion parameter θ assumed constant across origins s, destinations d, and occupations o, the scale parameters define fundamental comparative advantage of source countries across occupations in each destination.

We assume that migration is costly, where $c^s_{d,o}$ is the fraction of earnings lost for a worker from origin s to live in destination d and work in occupation o.[13] The migration cost $c^s_{d,o}$ captures all sources of migration frictions, including the monetary cost of moving to the destination (which likely varies by origin-destination pair but not by occupation), the cost of securing a visa for the destination (which because of rules for allocating visas likely varies by origin-destination and possibly by occupation), and the cost of finding a job once in the destination (which because of migration networks, and visa-based work restrictions likely varies by origin-destination and possibly by occupation).[14] We denote $\tau^s_{d,o} \equiv 1 - c^s_{d,o}$ as the fraction of earnings that a worker takes home.[15]

2.2 Labor allocation

Workers choose a country of residence and an occupation in order to maximize their perceived take-home wage, $\tau^s_{d,o} \cdot w_{d,o} \cdot a^{i,s}_{d,o}$, where $w_{d,o}$ is the wage per efficiency unit of labor in occupation o in destination d. Given the assumption of Fréchet distributed productivities, the fraction of s workers who live in country d and work in occupation o is

$$\Pi^s_{d,o} = \frac{\big(T^s_{d,o}\tau^s_{d,o}w_{d,o}\big)^\theta}{\sum_{d',o'}\big(T^s_{d',o'}\tau^s_{d',o'}w_{d',o'}\big)^\theta} \tag{1}$$

[12] Alternatively, we could allow Fréchet productivity draws to be correlated across occupations and destination countries, with differing cross-occupation and cross-destination correlations. Any such generalization of the Fréchet productivity distribution would produce an estimating equation that is isomorphic to ours, as long as we assume a common elasticity of substitution in labor supply at the occupational level.

[13] The assumption of multiplicative migration costs is common in the literature on the self-selection of immigrants (see, e.g., Borjas, 1987; Chiquiar and Hanson, 2005).

[14] On occupational licensing for foreign-born workers, see Han and Kleiner (2016); on occupational discrimination against foreign-born workers, see Oreopoulosa (2011).

[15] Because of occupational licensing, it is not necessarily true that $\tau^d_{d,o}$=1 for all o.

Conditional on having chosen to live in destination d, the fraction of workers from origin country s who work in occupation o is then

$$\Pi^s_{o|d} = \frac{\Pi^s_{d,o}}{\sum_o \Pi^s_{d,o}} = \frac{\left(T^s_{d,o}\tau^s_{d,o}w_{d,o}\right)^\theta}{\sum_{o'}\left(T^s_{d,o'}\tau^s_{d,o'}w_{d,o'}\right)^\theta} \tag{2}$$

In the empirical analysis, we study the allocation of college-educated workers from each origin country across occupations in two destination countries, the US and Canada.

We first use (2) to present visual evidence of comparative advantage by constructing a double difference, which is the log ratio workers born in origin s (e.g., college-educated men born in China) and workers born in destination d (e.g., college-educated men born in the US) who are employed in occupation o (e.g., computer programming) minus the equivalent log ratio of the employment shares for the two groups of workers in some base occupation o',

$$\log \frac{\Pi^s_{o|d}}{\Pi^d_{o|d}} \Big/ \frac{\Pi^s_{o'|d}}{\Pi^d_{o'|d}} = \theta \cdot \log \frac{T^s_{d,o}}{T^d_{d,o}} \Big/ \frac{T^s_{d,o'}}{T^d_{d,o'}} + \theta \cdot \log \frac{\tau^s_{d,o}}{\tau^d_{d,o}} \Big/ \frac{\tau^s_{d,o'}}{\tau^d_{d,o'}}. \tag{3}$$

The difference in the numerator cancels out the destination d wage per efficiency unit in occupation o, $W_{d,o}$; the difference in the denominator does the same for $W_{d,o'}$. The difference of the numerator from the denominator cancels out the corresponding denominator in (2), which summarizes employment opportunities for workers from s or d in destination d. In order for this comparison to be insensitive to the choice of base occupation o', we construct the denominator $\log \frac{\Pi^s_{o'|d}}{\Pi^d_{o'|d}}$ using the geometric mean over all occupations, such that,

$$\log \frac{\Pi^s_{o'|d}}{\Pi^d_{o'|d}} = \exp\left[\frac{1}{\text{dim}(o)} \sum_o \log \frac{\Pi^s_{o|d}}{\Pi^d_{o|d}}\right]. \tag{4}$$

We next use (2) to derive an estimating equation for the determination of the log share of workers born in source s who are living in d and working in occupation o,

$$\log \Pi^s_{o|d} = \theta \log T^s_{d,o} + \alpha_s + \alpha_o + \theta \log \tau^s_{d,o}. \tag{5}$$

The term, $\alpha_s \equiv -\log \sum_{o'}\left(T^s_{d,o'}\tau^s_{d,o'}W_{d,o'}\right)^\theta$, is source-country fixed effect that captures average employment opportunities for s workers in other occupations in d; note that this term will also absorb any migration costs that are specific to origin s and common across occupations in d (e.g., financial and psychic costs of moving, visa application costs). The term, $\alpha_o \equiv \theta \log W_{d,o}$, is an occupation fixed effect, which absorbs the average price per efficiency unit of labor in d for occupation o; this term also absorbs any occupational licensing costs specific to o that are common to workers regardless of their birth country. The final term, $\theta \log \tau^s_{d,o}$, captures bilateral frictions that affect migration and occupation choices. Below, we discuss how we specify the determinants of $T^s_{d,o}$ and $\tau^s_{d,o}$ in (5).

2.3 Sorting with Multi-Dimensional Skills

In standard applications of Fréchet-Roy, skill is one dimensional: a worker's productivity in each occupation is given by a scalar function of a Fréchet draw. In our application, we follow Lindenlaub (2017) and assume that worker skill is multi-dimensional.[16] Workers vary in their cognitive skill (i.e., the ability to excel in problem solving) and linguistic proficiency (i.e., facility in languages that are spoken in the US), where these skills affect task-specific productivity based on task-specific parameters. We model positive sorting on multi-dimensional skills to provide a theoretical interpretation of our empirical results.

To motivate our approach, consider empirical evidence on the labor market return to education across occupations differentiated by their task intensity. Using the Princeton Data Improvement Initiative survey, Autor and Handel (2013) show that college educated workers are much more likely to perform abstract tasks on the job than are non-college-educated workers.[17] This regularity is consistent with the positive sorting of workers across jobs, where educational attainment is the relevant skill for abstract tasks. Holding education constant, workers in more abstract-task-intensive jobs earn higher wages,[18] which is consistent with higher ability workers sorting into jobs more intensive in abstract tasks (i.e., positive sorting in unobserved skill). We interpret these results as supporting our assumption that workers with stronger cognitive skills will sort into jobs more intensive in cognitive tasks. Regarding language, Dustmann and Fabbri (2003) and Bleakley and Chin (2004) find that immigrants with better English language skills earn higher wages in the UK and US labor markets, respectively, while Peri and Sparber (2009) find that in the US less-educated immigrants from non-English speaking countries are less likely than US natives to be employed in communication-intensive occupations. These results are consistent with positive returns to ability in the native language and with positive sorting based on language ability.

To operationalize positive sorting in multi-dimensional skill in our context, we assume that group-level occupational efficiency $T^s_{d,o}$ (e.g., the capability of college graduates born in China in medical science) is an exponential function of skills,

$$T^s_{d,o} = \prod_{X_o \in \Omega} \exp\left(\beta_x X_o \mathrm{Cog}_s + \gamma_x X_o \mathrm{Ling}_s\right) \tag{6}$$

[16] Lindenlaub's model addresses two-side matching between firms and workers, where firms care about individual worker productivity. In her case, the matching function is characterized by a differential equation. In our model, sorting occurs on the labor supply side only. Implicitly, firms care not about individual worker productivity but the total efficiency units of labor they hire.

[17] In these data, abstract task intensity is the first principal component across four task measures: length of longest document typically read as part of the job, frequency of mathematics tasks involving high school or higher mathematics, frequency of problem-solving tasks requiring at least 30 minutes to find a good solution, and proportion of workday managing or supervising other workers.

[18] These results hold for abstract task intensity defined either in the PDII or O*NET.

where Cog_s is the cognitive skill of workers from country s (e.g., PISA math score) and $Ling_s$ is the linguistic aptitude of workers from s (e.g., proficiency in English). The vector $X_o \in \Omega = \{cog_o, com_o, rou_o, man_o\}$ contains occupation intensities for cognitive, communication, routine, and manual tasks, respectively.[19] The combination $\beta_x cog_o$ is the marginal productivity that converts cognitive skill into cognitive task output in occupation o, $\gamma_x com_o$ is the marginal productivity that converts linguistic skill into communication task output, and so forth. The aggregator in (6) is the basis for our empirical specification.

In equation (6), the variation in group-level occupational productivity $T^s_{d,o}$ is driven by differences in the cognitive and linguistic skills of workers. Because we lack country-level measures of routine and manual skills, we abstract from their role in occupational sorting.[20] Our empirical analysis is thus informative about sorting in two dimensions—cognitive and linguistic skill—and silent about sorting in other dimensions.

Unlike the case of uni-dimensional skill, workers in our context cannot be fully ranked by their skill levels. Leveraging results in Lindenlaub (2017), we state the conditions necessary for positive sorting of workers across occupations by skill in the case of two-dimensional skills (cognitive and linguistic skill). Let the occupational choice probability (i.e., the assignment function) be $\Pi(\mathrm{Cog}_s, \mathrm{Ling}_s, cog_o, com_o)$, which we assume is continuous and twice differentiable. Following Lindenlaub (2017), we have positive sorting in two-dimensional skills if there is positive sorting along each individual dimension of skill, and sorting within the "natural" dimension of each skill is more pronounced than across the "natural" dimensions of the skills. Stated formally, positive sorting requires,

(A) $\frac{\partial^2 \Pi(\mathrm{Cog}_s,\mathrm{Ling}_s,cog_o,com_o)}{\partial \mathrm{Cog}_s \, \partial cog_o} > 0$

(B) $\frac{\partial^2 \Pi(\mathrm{Cog}_s,\mathrm{Ling}_s,cog_o,com_o)}{\partial \mathrm{Ling}_s \, \partial com_o} > 0$

(C) $$\begin{vmatrix} \frac{\partial^2 \Pi(\mathrm{Cog}_s,\mathrm{Ling}_s,cog_o,com_o)}{\partial \mathrm{Cog}_s \, \partial cog_o} & \frac{\partial^2 \Pi(\mathrm{Cog}_s,\mathrm{Ling}_s,cog_o,com_o)}{\partial \mathrm{Cog}_s \, \partial com_o} \\ \frac{\partial^2 \Pi(\mathrm{Cog}_s,\mathrm{Ling}_s,cog_o,com_o)}{\partial \mathrm{Ling}_s \, \partial cog_o} & \frac{\partial^2 \Pi(\mathrm{Cog}_s,\mathrm{Ling}_s,cog_o,com_o)}{\partial \mathrm{Ling}_s \, \partial com_o} \end{vmatrix} > 0.$$

Lindenlaub (2017)'s concept of multi-dimensional sorting is based on matching on observables. In her case, there is a one-to-one matching between workers and occupations (or full specialization). Our model has (unobserved) idiosyncratic Fréchet distributed productivities for each national-origin group of college-educated labor. Our definition thus extends Lindenlaub (2017) to a probabilistic case (or partial specialization).

To interpret the above conditions, note that (A) and (B) require that there be positive sorting along each of the single natural skill-task dimensions: workers from countries with

[19] In our empirical analysis, we will assume that these intensities are the same in the US and Canada.

[20] Implicitly, we assume that the average skills of college-educated workers in routine and manual activities are equal across origin countries for immigration.

stronger cognitive skills are more likely to choose cognitive-task-intensive occupations (Cog_s-to-cog_o), and workers from countries with stronger linguistic skills (vis-á-vis the US) are more likely to choose communication-task-intensive occupations (Ling_s-to-com_o). These conditions follow the probabilistic version of uni-dimensional sorting defined in Costinot and Vogel (2015), in that they imply that the Monotone Likelihood Ratio Property holds in each skill dimension considered individually.[21] Condition (C) distinguishes the multi-dimensional and uni-dimensional cases. For positive sorting to occur, sorting *within* the natural dimensions (characterized by the product of diagonal elements) must be more pronounced than sorting *across* the natural dimensions (characterized by the product of off-diagonal elements).

Substituting equation (6) into (5), we have

$$\log \Pi^s_{d,o} = \theta \sum_{X \in \Omega} \Big[\beta_X \cdot \text{Cog}_s \cdot X_o + \gamma_X \cdot \text{Ling}_s \cdot X_o \Big] + \alpha_s + \alpha_o + \theta \log \tau^s_{d,o}. \tag{7}$$

In the empirical analysis, we will measure X_o using data on occupational task intensities, which allows us to estimate the β_X and γ_X vectors. The testable implications of positive sorting with two skill dimensions (cognitive and linguistic skills) are

(A′) $\beta_{cog} > 0$

(B′) $\gamma_{com} > 0$

(C′) $\begin{vmatrix} \beta_{cog} & \beta_{com} \\ \gamma_{cog} & \gamma_{com} \end{vmatrix} = \beta_{cog}\gamma_{com} - \beta_{com}\gamma_{cog} > 0.$

Since we do not observe the country-level skills that are relevant for routine and manual operations, we are unable to examine sorting within the natural dimension of these capabilities. Nevertheless, we add the interaction terms, Cog_s-to-rou_o, Cog_s-to-man_o, $Ling_s$-to-rou_o, and $Ling_s$-to-man_o, as regressors. Their inclusion allows us to control for sorting across the natural dimensions of these other skill types. Further interacting occupational task intensities with origin-country geographic distance to the US provides an additional set of controls for migration costs that may be specific to the occupation.

3 Data

Our goal is to estimate equation (7) for foreign-born college-educated workers in the US and Canada and to use the results to evaluate the importance of comparative advantage in foreign-born worker sorting across occupations. In this section, we discuss how we measure

[21] The implication of (A) is that for $\widetilde{\text{Cog}}_s > \text{Cog}_s$ and $\widetilde{cog}_o > cog_o$, then

$$\frac{\Pi(\widetilde{\text{Cog}}_s, \text{Ling}_s, \widetilde{cog}_o, com_o)}{\Pi(\widetilde{\text{Cog}}_s, \text{Ling}_s, cog_o, com_o)} > \frac{\Pi(\text{Cog}_s, \text{Ling}_s, \widetilde{cog}_o, com_o)}{\Pi(\text{Cog}_s, \text{Ling}_s, cog_o, com_o)}.$$

origin country cognitive and linguistic skills, occupational requirements in specific tasks, the allocation of workers across occupations, and other details. We then present graphic evidence on comparative advantage in immigrant sorting across occupations.

3.1 Origin Country Cognitive and Linguistic Skills

Following recent literature on the contribution of education to economic development (see, e.g., Hanushek and Woessmann, 2011; Woessmann, 2016), we measure cognitive skill in origin countries using PISA exam scores. PISA exams are international assessments of the scholastic performance of 15-year-old students in mathematics, science, and reading. They are widely used to measure national educational quality (e.g., Guiso et al., 2008; Fryer and Levitt, 2010; Bharadwaj et al., 2012). The exams, which are conducted by the Organization for Economic Cooperation and Development, have been administered every three years since 2000. Participation by countries has varied over time, rising from 41 countries in 2000 to 70 countries in 2009 and to 75 countries in 2015. By summarizing cross-country differences in academic performance, the exams provide information about the success of national educational institutions in training students to reason analytically.

PISA exam scores are strongly positively correlated across math, science, and reading assessments. In 2012, the correlation between country average math and science scores was 0.97, between average math and reading scores was 0.96, and between average science and reading scores was 0.98. These patterns suggest that the exams capture the general ability to solve problems, rather than aptitude in specific subjects. For this reason, we interpret PISA scores as a measure of cognitive skill. We use math scores in our analysis; results are unchanged when using science or reading scores, instead. Across schools, PISA scores are positively correlated with the amount of time that students spend in instruction, although this association is weaker in lower income countries (Lavy, 2015). Across countries, PISA scores are positively correlated with the rate of time preference and per capita GDP growth, but not with education spending per student (Hanushek et al., 2020). Within the US, origin-country PISA scores are positively correlated with the labor market returns to education when estimated separately for immigrants by country of birth (Schoellman, 2012).

It is also true that PISA exam scores are strongly persistent over time, as seen in Figure 1. The correlation between average math scores in 2000 and 2009 (40 countries) is 0.91, while that for 2009 and 2015 (62 countries) is 0.93. Such persistence suggests, not unreasonably, that national education quality changes slowly over time. We exploit this persistence to increase our sample size. Because not all countries participated in the tests in all years, we average exam scores over the 2000 to 2015 period, which produces a sample of 69 countries for which we have data on foreign-born workers in the US. If international migrants are positively selected in terms of education or other determinants of skill (Grogger and Hanson, 2011), then average scores in a country may not be indicative of the skills of workers who migrate abroad. To account for positive selection, we also report results using the 75^{th} and 90^{th} percentiles of PISA exam scores, averaged over the years 2009, 2012, and 2015, for which these moments are available. When using these measures, our sample is 61 countries.

There are well-known limitations of PISA exam scores. One is that students who participate in the tests may not be nationally representative. Although most countries administer PISA exams on a nationwide basis, others limit exams to wealthier or better-educated regions (Sands, 2017). For example, in the 2012 round China only tested students in Shanghai (though in the 2015 round it expanded tests to include the five largest provinces). Using higher moments of PISA exam scores to measure cognitive skill (e.g., scores at the 90^{th} percentile) may help attenuate the effects of cross-country differences in student testing protocols. A second limitation is that some origin countries, including India, do not participate in the PISA exams and therefore are excluded from the analysis.

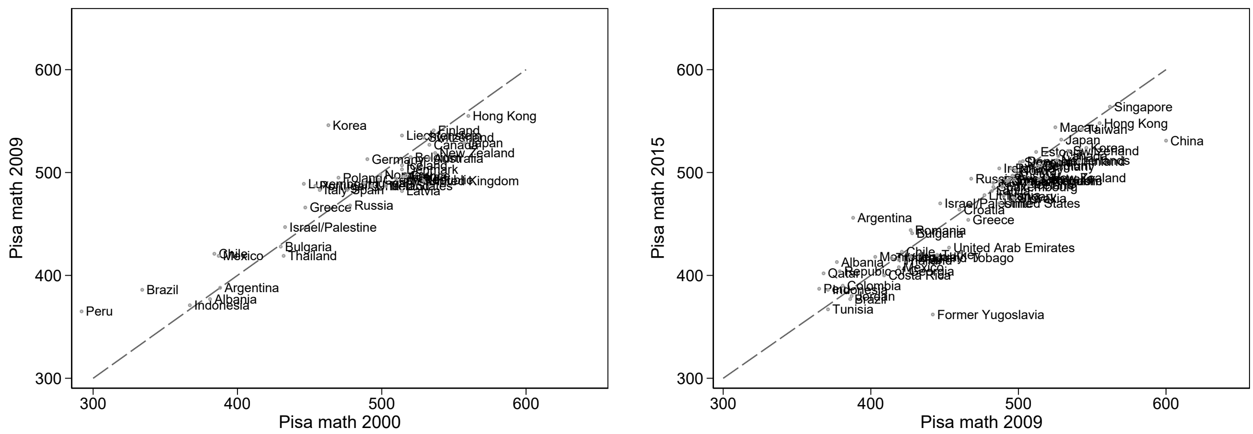


Figure 1: PISA math scores in 2000 and 2009 (left panel), 2009 and 2015 (right panel)

An alternative source for data on country education quality is the Trends in International Mathematics and Science Study (TIMSS). Because TIMSS only covers 50 countries and disproportionately represents the Middle East, it is less suitable for our analysis. PISA exams cover a larger number of countries and have stronger representation among countries in Latin American and East and Southeast Asia. Among the 42 countries for which both PISA and TIMSS math scores are available, the correlation in average scores is 0.83.[22]

To measure linguistic proximity between the US and Canada and origin countries for immigration, we use data from Melitz and Toubal (2014).[23] They construct linguistic proximity between countries using the Automatic Similarity Judgement Program (ASJP) from the Max Planck Institute, which transcribes 40 common words into phonetic script (ASJP code) for each major language and counts the number of phonetic changes that separate each word

[22]In previous work, Hanushek and Kimko (2000) combine multiple tests of student achievement in mathematics and science conducted over 1960 to 1990. Their combined sample has 39 countries.

[23]See Isphording and Otten (2014) on how greater linguistic distance and impedes language acquisition among immigrants in Germany and the US.

for a large number of language pairs. For instance, Table 1 shows that whereas one phonetic change separates the English and German words for "you", two phonetic changes separate the English and Chinese version of the word; similarly, whereas three changes separate the English and German words for "name", five changes separate the English and Chinese versions of the word. The average number of changes across the 40 words is a measure of linguistic distance between two languages. To calculate linguistic proximity between a pair of countries, Melitz and Toubal (2014) use the population-weighted average of language-to-language linguistic proximity, where the weights are the products of population shares for the two most commonly spoken languages in each country. The appeal of this approach for measuring linguistic proximity, as opposed to commonly used measures such whether two countries share an official language, is that it captures differences between countries in how words are pronounced. Such differences are likely to be important when it comes to performing communication-intensive tasks on the job.

Table 1: Examples of ASJP Codes for English, German, and Chinese

Word	English ASJP code	Germany ASJP code	linguistic distance	Chinese ASJP code	linguistic distance
you	yu	du	1	ni	2
name	neim	nome	3	minci	5

Linguistic distance is the minimum number of changes in ASJP codes needed to translate words between languages (e.g., to move from "neim" (English ASJP code for "name") to "nome" (German ASJP code), one needs three steps: change the letter "e" to "o", delete the letter "i", and add letter "e" to the end).

As a proxy for migration costs, we use bilateral geographic distance from CEPII.[24] We interact bilateral distance with the full set of occupational task intensities, to account for differential sorting of migrants across occupations based on physical migration costs. Origin country fixed effects in the regression helpfully absorb bilateral migration costs that are common across college-educated workers from a given birth country.

3.2 Occupation Task Intensity

Recent literature on how technology, globalization, and other shocks affect labor-market outcomes differentiate occupations according to their task intensities and skill requirements (e.g., Autor et al., 2003; Goos and Manning, 2007; Autor and Dorn, 2013). If we know which tasks are performed by which types of workers, we can evaluate how changes in the demand for tasks or the supply of skills affects wages and employment across occupations. To

[24]Our analysis uses the CEPII variable $distwt$, which measures bilateral distance weighted by countries' internal population distributions and cross-city distances.

measure task requirements, researchers frequently use the US Department of Labor's Dictionary of Occupational Titles (DOT) or its successor, the Occupational Information Network (O*NET). Because in both the DOT and O*NET the scale of intensity for individual tasks is unknown (Acemoglu and Autor, 2011), it is unclear which data source is a better guide for measuring occupational tasks requirements. To compensate, we utilize both DOT and O*NET-based measures of task intensity in our analysis.

Using the fourth edition of the DOT, we measure cognitive, routine, manual, and communication task requirements by occupation. The first three measures, which we take from Autor, Levy and Murnane (2003), identify cognitive (or abstract) tasks as those involving abstract problem solving, and creative, organizational, and managerial tasks; routine tasks as routine, codifiable cognitive, and manual operations that follow explicit procedures; and manual tasks as those requiring non-routine manual operations that in turn require physical adaptability. We denote these measures as cog_o, rou_o, and man_o. Following Shu et al. (1996), we measure communication task intensity in the DOT using the variable "talk," which indicates demands for listening and speaking on the job; we denote it as com_o.[25]

Our O*NET measures of task intensity follow Deming (2017). Cognitive task intensity is the average of mathematical reasoning ability, mathematics knowledge, and mathematics skill; routine task intensity is the average of the degree of automation and importance of repeating the same tasks; manual task intensity is the average of assisting and caring for others, and service orientation; and communication task intensity is the average of social perceptiveness, coordination, persuasion, and negotiation. Note that whereas the DOT measure of communication task intensity emphasizes speaking and listening on the job, the O*NET measure is a broader concept that emphasizes the use of social skills.

Since neither DOT nor O*NET variables have a natural scale, we compute the percentile ranking of each task using IPUMS occ1990 codes and weight by hours worked to construct percentile measures for our 29 occupational categories. Details are in Appendix A. DOT and O*NET measures of task intensity are high correlated for cognitive and communication tasks (0.73 and 0.77, respectively) and less highly correlated for routine tasks and manual tasks (0.20 and 0.22, respectively). We obtain similar results for the two sets of measures.

3.3 Employment of Foreign-born Workers in the US and Canada

Our analysis uses data on the employment of foreign-born workers in two destination countries, the US and Canada. For each country, we restrict the sample to prime-age males (25 to 54 years old), who have at least four years of college education, who earn positive wages, and who do not reside in group quarters. We restrict the analysis to men because for many origin

[25]First, we match the 3,886 DOT occupations to the 327 IPUMS OCC1990 occupations in our data (using data for 1970 as a crosswalk). Second, we aggregate occupations that are similar in their task contents to our 29 aggregate occupational categories. We proceed similarly when using O*NET data.

countries the number of foreign-born college-educated women with a job is small, leading to a preponderance of zero cells at the occupation level.

For US employment data, we use the three-year American Community Survey (ACS) sample for 2011 to 2013. Our choice of this time period is to match the data we have available for Canada. We measure employment as the share of total hours worked for a given origin-country group in a given occupation.[26] We group detailed occupations into 29 categories based on similarity in occupational task requirements. Table A.2 lists the occupation categories and displays their DOT and O*NET task intensities.

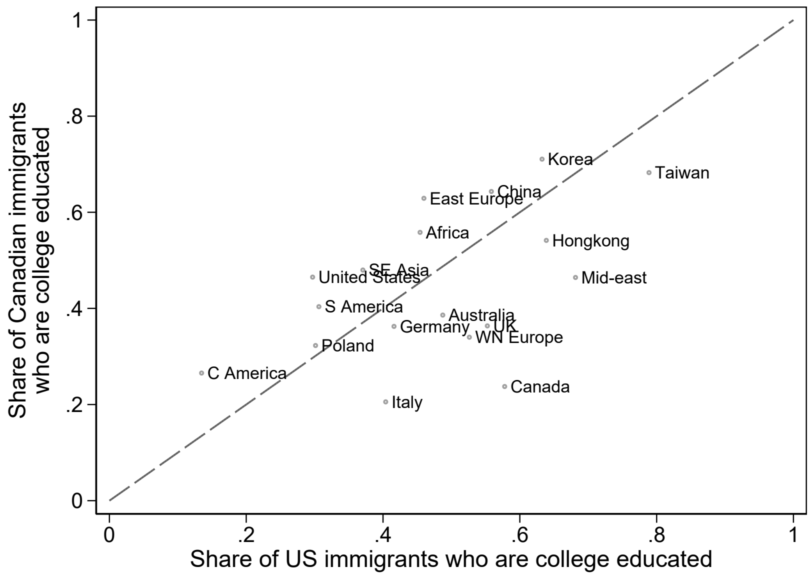


Figure 2: Share of prime-age foreign-born male workers with at least a college education in Canada (2011) and the US (2011-2013), by origin-country region

Our data for Canada are from the 2011 National Household Survey (NHS). Although micro data were unavailable, we obtained from Statistics Canada tabulations of total hours worked by occupation and country of birth for college-educated men 25 to 54 years of age. Because for some origin countries the Canadian Census reports immigrant country of birth at a geographically aggregated level, we cluster origin countries into 22 regions. Canada classifies occupations according to a four-digit National Occupational Classification (NOC) code. We aggregate 4-digit NOC codes to match as closely as possible our 29 US occupation categories.[27] We then match DOT and O*NET task variables to these occupations.

[26]The weight of each individual in the count of hours worked is given by,

$$\frac{\text{Census sampling weight} \times \text{weeks worked} \times \text{usual hours per week}}{2000}.$$

where the value for weeks worked is the mid point of the interval reported in IPUMS.

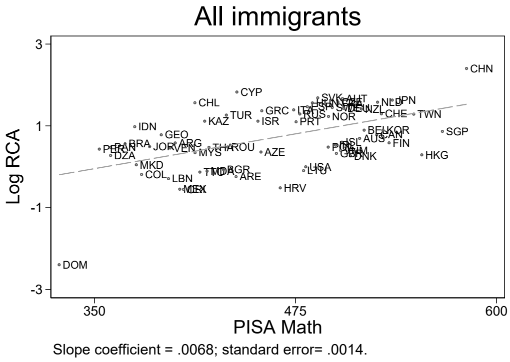
[27]Because NOC codes pool construction workers, machine operators, and transportation workers in a single category (which are separate occupations in US data), we have 27 Canadian occupational categories.

To evaluate the degree of similarly in education levels of immigrants in the US and Canada, Figure 2 plots the shares of adult male immigrants (25 to 54 years old) with at least a college education by region of birth for the two countries in 2010. Most origin regions lie close to the 45-degree line, indicating similar patterns of college attainment among immigrants by region of birth in the two destinations. Using data for 1980, Borjas (1993) documents a corresponding similarity in the educational attainment of immigrants by birth region in the US and Canada. Despite very different immigration policy regimes, it appears that the US and Canada attract individuals with concordant levels of educational attainment by origin country and have done so for the last several decades.

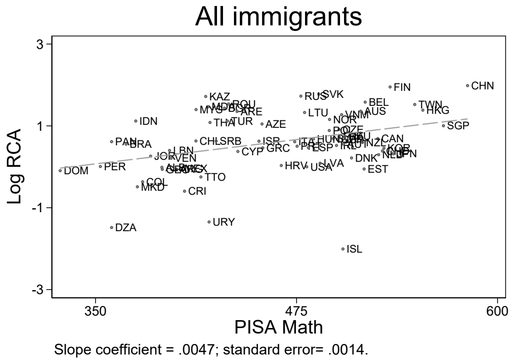
3.4 Preliminary Evidence

Before turning to the regression analysis, we present graphic evidence on occupational comparative advantage. We plot occupational specialization (relative to the average occupation) by origin country for college-educated immigrants in the US, as shown in equation (4), against origin-country average PISA math scores. Figure 3 includes four occupations: two are among the top-ranked in terms of DOT or O*NET cognitive task intensity— scientists and mathematicians and computer software developers—and two are ranked toward the bottom—salespersons and high-skill clerical workers. For the cognitive-task-intensive occupations, we see a strong positive correlation between immigrant specialization and origin-country PISA scores: foreign-born workers from countries whose students score more highly on PISA exams are more likely to specialize in occupations that are intensive in cognitive reasoning (precisely estimated slope coefficients of 0.68 in panel (a) and 0.47 in panel (b)). We see no such correlation between PISA scores and immigrant specialization in sales or clerical work (imprecisely estimated slope coefficients of 0.18 in panel (c) and -0.15 in panel (d)). The visual evidence in Figure 3 is consistent with foreign-born workers from countries whose educational institutions produce students who score highly on international assessments concentrating more strongly in jobs that require more abstract problem solving.

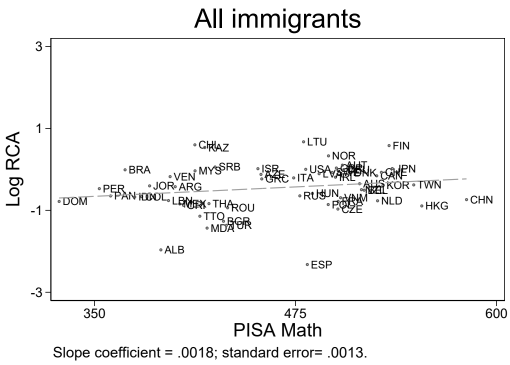
There is, however, a potential source of ambiguity in the results in Figure 3. Because the population of foreign-born workers in the US combines individuals who were educated in their country of birth and individuals who migrated to the US at a young age and attended US schools, the evidence on the importance of being educated in the origin country for occupational specialization is not dispositive. To help resolve this ambiguity, in Figure 4 we separate foreign-born workers into those who arrived in the US at age 18 or older (row one), and therefore were likely to have completed their K-12 education in their country of birth, and those who arrived in the US at age 17 or younger (row two), and therefore were likely to have completed part or all of their K-12 schooling in the US. Mathematicians and scientists appear in column one; computer software developers appear in column two. The evidence of comparative advantage in Figure 3 is clearly driven by immigrants who arrived in the US at age 18 or older. For this group, we see a strong positive correlation between origin-country PISA scores and immigrant specialization in math and science and computer software (precisely estimated slope coefficients of 0.71 in panel (a) and 0.56 in panel (b)). For immigrants who arrived in the US at aged 17 or younger the correlations are much weaker (imprecisely estimated slope coefficients of 0.03 in panel (c) and 0.13 in panel (d)).



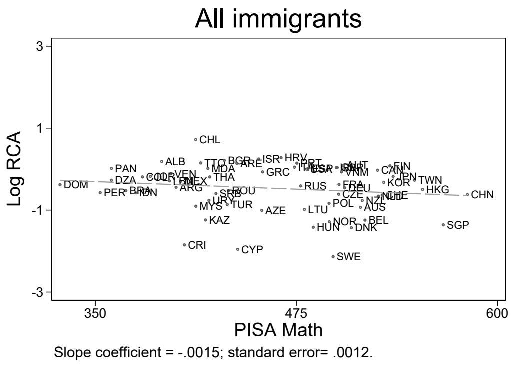
(a) Mathematicians and scientists



(b) Computer software developers



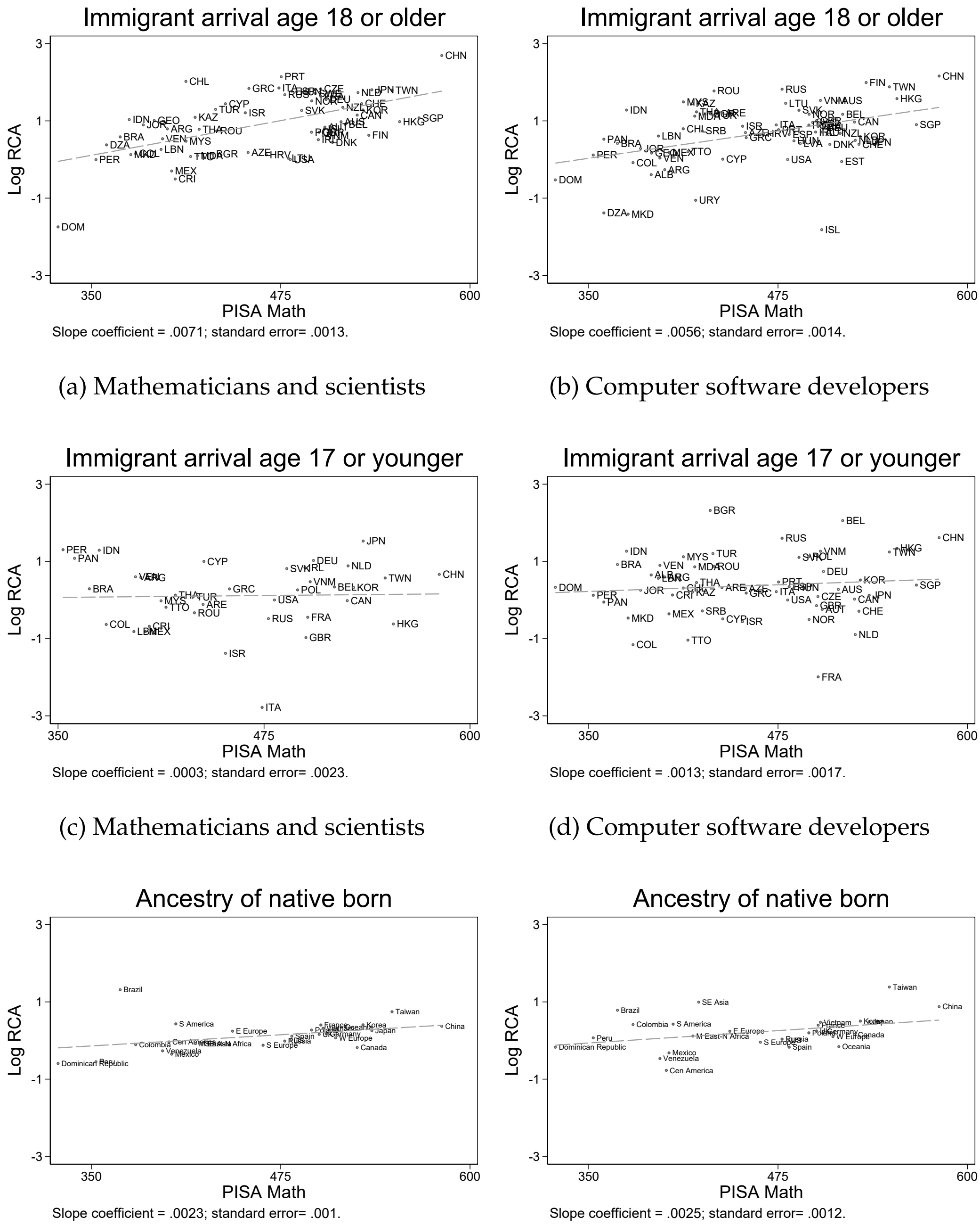
(c) Salespersons



(d) Clerical workers, high-skill

Figure 3: Occupational specialization ($\log \frac{\Pi^s_{o|d}}{\Pi^s_{o'|d}} \Big/ \frac{\Pi^{s'}_{o|d}}{\Pi^{s'}_{o'|d}}$) and PISA math score by immigrant origin-country, various occupations

The distinction in specialization patterns among immigrants according to their age of arrival in the US is important because it helps neutralize a confounding force for occupational specialization. If some cultures accord higher status to individuals who perform well in school (e.g., Figlio et al., 2019; Hanushek and Kimko, 2000), then the correlation between country test scores and immigrant specialization patterns could reflect these cultural or social values, rather than the actual skills imparted by origin-country schools. The evidence in Figure 4 is contrary to such a cultural values hypothesis.



(a) Mathematicians and scientists

(b) Computer software developers

(c) Mathematicians and scientists

(d) Computer software developers

(e) Mathematicians and scientists

(f) Computer software developers

Figure 4: Occupational specialization ($\log \frac{\Pi^s_{o|d}}{\Pi^s_{o'|d}} \Big/ \frac{\Pi^{s'}_{o|d}}{\Pi^{s'}_{o'|d}}$) and PISA math score by immigrant origin-country, age of arrival in the US, and region of ancestry

To evaluate cultural factors further, in row three of Figure 4 we present results for US-born individuals categorized by their country or region of ancestry. If perceptions of social status are transmitted culturally, then the educational and occupational choices of individuals may persist across generations and follow patterns associated with the origin countries of their elders (e.g., Figlio and Özek, 2020). In the ACS, ancestry is an individual's self-reported country of ancestry (e.g., the birth country of one's parents, grandparents, or great grandparents).[28] Since ancestry is sometimes identified as a geographic region instead of a country, we are forced to use more aggregate country groupings than in the previous analysis for foreign-born individuals. We have 25 to 28 ancestral regions (depending on the test-score measure used), which are comprised of 16 to 19 individual countries and eight aggregate regions.[29] In row three of Figure 4, we see weak correlations (imprecisely estimated slope coefficients of -0.15 in panel (e), 0.03 in panel (f)) between occupational specialization of individuals classified by their ancestral countries and PISA scores in these countries. These results are further evidence against cultural values explaining occupational specialization.

4 Empirical Results

Our regression specification, based on equation (7), is for the log share of hours worked in occupation o by prime-age college-educated workers born in country s and residing destination country d (the US or Canada). It is given by,

$$\log \Pi^s_{d,o} = \sum_{X_o \in \Omega} \left[\beta_X \cdot \text{Cog}_s \cdot X_o + \gamma_X \cdot \text{Ling}_s \cdot X_o + \phi_X \cdot Dist_s \cdot X_o \right] + \alpha_s + \alpha_o + \varepsilon^s_{d,o}.^{30} \quad (8)$$

We interact the country s PISA math score (Cog_s), linguistic proximity to the destination (Ling_s), and the log of geographic distance to the destination (Dist_s) with DOT or O*NET measures of occupational intensities in cognitive, communication, routine, and manual tasks (X_o), where α_s is a fixed effect for country s (absorbing employment alternatives for country s workers and migration costs from s to the destination that are common across occupations) and α_o is a fixed effect for occupation o (absorbing the occupation wage per efficiency unit in the destination and occupational credentialing costs that are common across workers regardless of their nationality). We cluster standard errors by country of origin.

4.1 Occupational Specialization in the US

4.1.1 Benchmark Results

In Table 2, we present results using US data on occupational specialization and DOT measures of task intensity. We choose DOT occupation measures for our baseline because its communication measure captures the intensity of speaking and listening on the job, whereas the O*NET measure is more properly thought of as capturing intensity in the use of social skills (Deming, 2017). We limit the reported coefficients estimates to those that allow us to

[28] Our partition of ancestry groups is based the ACS variable (ANCESTR1), which records a respondent's first response for ancestry or ethnic origin.

[29] The individual countries are Brazil, Canada, China, Colombia, the Dominican Republic, France, Germany, Japan, Korea, Mexico, Peru, Poland, Russia, Spain, Taiwan, the US, Venezuela, and Vietnam; the aggregate regions are Central America and the Caribbean, Eastern Europe, the Middle East and North Africa, Oceania, South America, Southern Europe, Southeast Asia, and other Western Europe.

[30] β_X and γ_X differ from those in (6) by a common multiplier θ. We normalize $\theta = 1$ to simplify notation.

evaluate positive worker sorting in cognitive and linguistic skills, as summarized in conditions (A′) to (C′). These results are for the interactions of PISA math scores (Cog$_s$) and linguistic proximity (Ling$_s$) with occupational intensities in cognitive and communication tasks. Results for our hypothesis tests on positive sorting are in Table 3; regression results for the complete set of interactions appear in Appendix Table C.1.

Table 2: OLS results for $\log \Pi^s_{d,o}$ (log share of hours worked) for US foreign-born Workers

	(1) Average math	(2) 75ptl math	(3) 90ptl math
	All foreign-born workers		
$Cog_s \times cog_o$	8.663	8.580	10.46
	(2.641)	(2.215)	(2.155)
$Cog_s \times com_o$	1.975	1.620	2.338
	(2.154)	(2.345)	(2.213)
$Ling_s \times cog_o$	0.752	0.738	0.745
	(0.379)	(0.427)	(0.402)
$Ling_s \times com_o$	0.956	0.966	0.900
	(0.446)	(0.454)	(0.444)
Observations	1809	1597	1751
Adjusted R^2	0.311	0.318	0.313
Number of countries	69	61	67
	Summary statistics for $\log \Pi^s_{d,o}$		
Mean	Standard deviation	25ptl	75ptl
-0.18	0.86	-0.69	0.35

Notes: This table reports selected OLS estimation results for equation (8). The sample covers 29 occupations and 61 to 69 origin countries. Hours worked are for male workers who are 25 to 54 years old, have at least four years of college education, earn positive wages, and do not reside in group quarters. Occupation and country of origin fixed effects are included in all regressions. Standard errors are clustered by origin country and reported in parentheses. See Appendix Table C.1 for complete regression results.

In column (1), we measure PISA math scores using the average for a country, while in columns (2) and (3) we measure scores using the 75^{th} and 90^{th} percentiles. The choice of test-score moment determines the number of countries for which we have data. To evaluate sorting within the natural dimension of cognitive skill, consider the coefficient estimate for the interaction of the PISA math score and cognitive task intensity, which is positive and highly precisely estimated ($\widehat{\beta}_{cog}$ = 8.7, t-value = 3.28). This interaction indicates that, consistent with Figure 3, workers from countries that score more highly on international assessments specialize more strongly in occupations that are more intensive in cognitive reasoning. To interpret the coefficient estimate, compare countries at the 25^{th} and 75^{th} percentiles of PISA math scores internationally. The higher scoring country would have a 86.7% higher share (1.01 of a standard deviation) of its workers in the US employed in management and finance, the occupation at the 75^{th} percentile of cognitive task intensity.[31]

Turning to linguistic skill, the positive and precisely estimated coefficient ($\widehat{\gamma}_{com}$ = 0.96, t-value = 2.14) on the interaction of linguistic proximity and communication task intensity indicates that workers from countries that are more linguistically similar to the US concentrate more heavily in occupations that are more intensive in communication on the job. Comparing countries at the 25^{th} and 75^{th} percentiles of linguistic similarity to the US, the more similar country would have a 20.8% higher share (0.24 of a standard deviation) of its US-based workers employed in executive management, which is the occupation at the 75^{th} percentile of communication intensity.[32] These first two sets of results are consistent with positive sorting of workers within the natural dimensions of skill.[33]

To evaluate results using other moments of the PISA test score distribution, Table 2 shows little change when we replace the average PISA score with the score at the 75^{th} (column 2) or 90^{th} (column 3) percentile. The similarity in results for alternative test-score moments suggests that the relevant difference in PISA scores for occupational sorting is the rightward or leftward shift of the score distribution for any individual country relative to other countries, rather than differences in score variance across countries. In Appendix Table C.3, we use O*NET measures of occupational task intensity in place of the DOT. There is a positive and significant interaction between PISA math scores and O*NET cognitive task intensity. With O*NET measures of social skill intensity (in place of DOT measures of communication intensity), the interaction with linguistic proximity to the US is imprecisely estimated.[34]

[31]The value of 86.7% is calculated as $\left(\widehat{\beta}_{cog} \cdot cog_o + \widehat{\beta}_{rou} \cdot rou_o + \widehat{\beta}_{man} \cdot man_o + \widehat{\beta}_{com} \cdot com_o\right) \times \left(\text{Cog}_s^{75} - \text{Cog}_s^{25}\right)$, where the $\widehat{\beta}$'s are estimated coefficients, Cog_s^{75} and Cog_s^{25} are the 25^{th} and 75^{th} percentiles of PISA scores, and DOT task intensities are for the management and finance occupation.

[32]The value of 20.8% is calculated as $\left(\widehat{\gamma}_{cog} \cdot cog_o + \widehat{\gamma}_{rou} \cdot rou_o + \widehat{\gamma}_{man} \cdot man_o + \widehat{\gamma}_{com} \cdot com_o\right) \times \left(\text{Ling}_s^{75} - \text{Ling}_s^{25}\right)$, where the $\widehat{\gamma}$'s are estimated coefficients, Ling_s^{75} and Ling_s^{25} are the 25^{th} and 75^{th} percentiles of linguistic proximity, and the DOT task intensities are for the executive management occupation.

[33]The OLS results in Table 2 omit observations for which country of birth-to-occupation matching is not observed in our 3% ACS sample. This omission may lead OLS results to underestimate true coefficient magnitudes if the missing countries have a weak comparative advantage in the occupations for which we observe no matching. In Appendix Table C.2, we re-estimate the regressions in Table 2 using Poisson pseudo maximum likelihood (PPML) (Silva and Tenreyro, 2006). The coefficient estimates are somewhat larger than in Table 2. Coefficient signs and patterns of significance are preserved, except for the linguistic proximity-communication intensity interaction, whose coefficient magnitudes are unchanged but whose standard errors are larger.

[34]This imprecision disappears when we restrict the sample to immigrants who arrived in the US at age 18 or older, as seen in Appendix Table C.6.

4.1.2 Sorting on Multi-Dimensional Skill

To assess the strength of positive sorting for the two skill dimensions together, we test condition (C′), alone, and conditions (A′), (B′), and (C′), jointly, as shown in Table 3. Positive sorting in two dimensions requires that $\beta_{cog}\gamma_{com} - \beta_{com}\gamma_{cog} > 0$, or that the product of sorting within natural skill dimensions (the diagonal interaction terms) exceeds the product of sorting across natural skill dimensions (the off-diagonal interaction terms). The sign condition for this determinant is satisfied, but, given imprecise estimates of the off-diagonal interaction terms, we are underpowered and fail to reject the determinant equaling zero at conventional significance levels (p-value = 0.12 in column 1).[35] When we test the joint hypothesis that (A′), (B′) and (C′) fail (i.e., that β_{cog} = γ_{com} = $\beta_{cog}\gamma_{com}$ -$\beta_{com}\gamma_{cog}$ = 0) we reject at any significance level. See Appendix B for details. We interpret these results as supporting the positive sorting of workers across occupations in two dimensions of skill.

Table 3: Tests of positive sorting: Conditions (A′), (B′), and (C′)

Sample: All immigrants			
Test of (C′)			
Test statistic	1.18	P-value	0.12
Joint test of (A′), (B′), and (C′)			
$\chi^2(3)$	113.1	P-value	0.00
Sample: Immigrants arriving in US age 18 or older			
Test of (C′)			
Test statistic	1.47	P-value	0.07
Joint Test of (A′), (B′), and (C′)			
$\chi^2(3)$	216.1	P-value	0.00

Notes: When testing (C′), we test against the alternative hypothesis $\beta_{cog}\gamma_{com} - \beta_{com}\gamma_{cog} > 0$; p-values are calculated based on a one-tailed test. When jointly testing (A′), (B′), and (C′), we test against the alternative hypothesis $\beta_{cog} > 0$, $\gamma_{com} > 0$, and $\beta_{cog}\gamma_{com} - \beta_{com}\gamma_{cog} > 0$; p-values are calculated based on a one-tailed chi-square distribution with three degrees of freedom. All results are based on regressions using the mean value of PISA scores from column 1 of Table 2.

Evaluating other interaction terms, as shown in Appendix Table C.1, those that are pre-

[35] Precision is greater when we restrict the sample to immigrants who arrived in the US at age 18 or older, as seen in the bottom panel of Table 3 and discussed below.

cisely estimated are a negative coefficient for the PISA score-manual task intensity interaction (indicating, intuitively, that individuals from high-test-score countries are relatively unlikely to select into manual-intensive occupations) and a negative coefficient on the geographic distance-manual task intensive interaction (indicating that workers from countries closer to the US are more likely to select into manual jobs). These results appear to be driven by differences between countries in Latin America and countries in other origin regions for US immigration. Latin America has among the lowest PISA scores for countries in our sample and comprises the countries (other than Canada) that are physically closest to the US, thereby allowing for more undocumented and low-skilled immigration (Grogger and Hanson, 2011). It thus appears that migrants from nearby countries with relative weak cognitive training are disproportionately likely to work in manual occupations.

4.1.3 Interpreting the Results

To aide in interpreting the results more fully, we calculate for each occupation the implied marginal change in occupational sorting (as measured by $\log \Pi^s_{d,o}$) in response to a one standard deviation change in the PISA math score or linguistic proximity to the US. The coefficients we use for PISA math scores are $\widehat{\beta}_{cog}$ and $\widehat{\beta}_{man}$, which correspond to the interaction terms $Cog_s \times cog_o$ (PISA score-cognitive task intensity) and $Cog_s \times man_o$ (PISA score-manual task intensity), respectively. For linguistic proximity, we use the coefficients $\widehat{\gamma}_{cog}$ and $\widehat{\gamma}_{com}$, which correspond to the interaction terms $Ling_s \times cog_o$ (linguisitic proximity-cognitive task intensity) and $Ling_s \times com_o$ (linguistic proximity-communication intensity), respectively. We limit the calculations to these coefficients because they are the ones with precisely estimated interactions between task intensities and PISA scores or linguistic proximity. All coefficients are from column (1) of Appendix Table C.1.

Table 4 reports the marginal changes in $\log \Pi^s_{d,o}$ for the five most positively and five most negatively affected occupations. Given the presence of occupation fixed effects in the underlying regression, these marginal changes are only meaningful with respect to some base occupation, which we select to be administrative managers given its place roughly in the middle of occupations in terms of cognitive and communication task intensities (see Appendix Table A.2). The values in parentheses are bootstrapped standard errors, calculated by drawing 200 independent samples (where bootstrapped samples are drawn from clustered country-occupation pairs). We estimate (8) for each sample, which provides 200 estimates of $\left(\widehat{\beta}_{cog} \times cog_o + \widehat{\beta}_{man} \times man_o\right) \times Std(Cog_s)$, where $Std(Cog_s)$ is the standard deviation of PISA math score, and which we evaluate relative to the administrative manager occupation. We calculate impacts for the increase in linguistic proximity analogously.

Table 4: Changes in $\log \Pi^s_{d,o}$ in response to one-standard-deviation increase in PISA math score and linguistic proximity to the US (relative to the administrative manager occupation)

Panel I: One standard deviation increase in PISA math score			
Occupations with the largest increase in $\log \Pi^s_{d,o}$		Occupations with the largest decrease in $\log \Pi^s_{d,o}$	
Accountants	0.122 (0.032)	Transportation workers	-0.437 (0.071)
Financial managers	0.112 (0.028)	Machine operators	-0.313 (0.059)
Post-secondary teachers	0.083 (0.018)	Middle-skill services	-0.297 (0.047)
Finance, insurance	0.071 (0.028)	Low-skill services	-0.278 (0.054)
Computer software developers	0.059 (0.029)	Mechanics & repairers	-0.266 (0.042)
Panel II: One standard deviation increase in linguistic proximity to US			
Occupations with the largest increase in $\log \Pi^s_{d,o}$		Occupations with the largest decrease in $\log \Pi^s_{d,o}$	
Primary, secondary teachers	0.064 (0.032)	Machine operators	-0.211 (0.072)
Post-secondary teachers	0.058 (0.029)	Transportation workers	-0.193 (0.062)
Other managers	0.028 (0.017)	Mechanics & repairers	-0.127 (0.044)
Health professionals	0.024 (0.011)	Low-skill clerical workers	-0.119 (0.041)
Lawyers	0.011 (0.029)	Construction	-0.109 (0.041)

Notes: The marginal changes are calculated using the expressions $(\widehat{\beta}_{cog} \times cog_o + \widehat{\beta}_{man} \times man_o) \times Std(Cog_s)$ in Panel A, and $(\widehat{\gamma}_{cog} \times cog_o + \widehat{\gamma}_{com} \times com_o) \times Std(Ling_s)$ in Panel B, where $Std(Cog_s)$ and $Std(Ling_s)$ are, respectively, the standard deviations of PISA math score and linguistic proximity. The standard errors in parentheses are estimated from 200 independent bootstrapped samples drawn from clustered country-occupation pairs.

Intuitively, an increase in the PISA math score increases the probability of working in cognitive-task-intensive occupations, while reducing the likelihood of employment in manual-task-intensive occupations. Relative employment shares increase by 12.2 log points for accountants and 11.2 log points for financial managers, while they decline by 43.7 log points for transportation workers and 31.3 log points for machine operators. An increase in linguistic proximity raises relative employment shares among primary and secondary teachers by 6.4 log points and post-secondary teachers by 5.8 log points—two occupations that are obviously intensive in communication in English—while reducing employment shares among machine operators and transportation workers by 21.1 and 19.3 log points, respectively. Three manual occupations—transportation, machine operation, and mechanics and repairers—see among the largest relative employment declines in response to either shock.

4.2 Occupational Specialization in Canada

Our results for the US are consistent with comparative advantage by origin country playing a significant role in the occupational choice of immigrant workers in the US. We see clear evidence of positive sorting in the dimensions of cognitive and linguistic skills. However, because the US immigration system is comprised of a complex set of preferences that govern which types of individuals are admitted, we cannot be sure that our results are not somehow the byproduct of differential national selection built into US visa policies.

Table 5: OLS regression results for $\log \Pi^s_{d,o}$, Canadian immigrants

	(1) Average math	(2) 75ptl math	(3) 90ptl math
$Cog_s \times cog_o$	9.241	8.573	8.752
	(3.147)	(2.936)	(2.891)
$Cog_s \times com_o$	0.364	0.649	0.818
	(4.045)	(3.802)	(3.939)
$Ling_s \times cog_o$	3.205	3.192	3.154
	(0.660)	(0.651)	(0.648)
$Ling_s \times com_o$	-0.147	-0.161	-0.171
	(0.706)	(0.711)	(0.721)
Observations	555	555	555
Adjusted R^2	0.717	0.718	0.719
Number of countries	22	22	22
	Summary statistics for $\log \Pi^s_{d,o}$		
Mean	Standard deviation	25ptl	75ptl
-0.26	0.90	-0.78	0.38

Notes: Regression units are 27 occupations × 22 aggregate regions. The sample is restricted to workers who are male and 25 to 54 years old, have a college education or above, and earn positive wages. Occupation and region-specific fixed effects are included in all regressions. Standard errors are clustered at the region level and reported in parentheses.

To help isolate the role of comparative advantage in immigrant sorting, we present results parallel to those in Section 4.1 for immigrants in Canada. In 1967, Canada became the first country to adopt a point system for awarding immigration visas. In its general outlines, Canada's system awards higher points, and therefore a higher likelihood of obtaining an immigration visa, to immigrants who are younger, more educated, have work experience, and can speak English and (or) French (Antecol et al., 2003; Schoellman, 2012). These criteria broadly favor workers with more human capital, rather than workers in specific occupations. Such a system is useful for evaluating the role of comparative advantage in occupational selection for prime-age, college-educated, foreign-born workers.

Table 5 reports our results for Canada, which like our US results cover foreign-born males who are 25 to 54 years old, have a college education or higher degree, and earn positive wages.[36] Because in some cases the Canadian Census reports the origin for immigrants at the region rather than country level, our data span 22 origin regions.[37] The positive and highly precisely estimated coefficient for the interaction of the PISA math score and cognitive task intensity ($\widehat{\beta}_{cog}$ = 9.2, t-value = 2.93) indicates that in Canada, as in the US, workers from regions that score more highly on international assessments specialize more strongly in occupations that are more intensive in cognitive tasks. Comparing regions at the 25^{th} and 75^{th} percentiles of PISA math scores, the higher scoring region would have a 67.1% higher share ($\widehat{\gamma}_{com}$ = 0.75 of a standard deviation) of its workers in Canada employed in management and finance (the occupation at the 75^{th} percentile of cognitive task intensity). These impacts are modestly smaller than those for the US.

In contrast to results for the US, the interaction of linguistic proximity and communication task intensity in the Canadian sample is small, negative, and highly imprecisely estimated (-0.15, t-value = 0.21). One explanation for this null result is that because Canada heavily rewards English or French speaking ability in granting visas, average linguistic proximity to Canada in an origin country may have little predictive power for the linguistic skills of the immigrants who ultimately qualify for admission to Canada and therefore may have little relevance for immigrant occupational sorting. Canada requires a minimum of 67 points to be eligible for a skilled worker immigration visa. An applicant may receive as many as 28 points toward this goal, based on language ability alone. For context, the maximum an applicant may receive based on educational attainment is 25 points.[38]

We interpret the results for Canada as confirming the role of comparative advantage in occupational selection. As in the US, there is positive sorting of workers in cognitive skill. Our results for the US do not appear to be purely a byproduct of US immigration policy.

4.3 Separating Immigrants by Age of Arrival in the US

Next, we group immigrants by their age of arrival in the US, which separates them based on where their K-12 education occurred. Doing so isolates the importance of being exposed to educational institutions in an immigrant's birth country, as opposed to being exposed to birth-country cultural values only (by virtue of having parents from a given origin). We focus on results using average PISA scores; results using other moments are very similar.

In panel I of Table 6, we limit the sample to immigrants who arrived in the US at age 18 or older and therefore completed their K-12 education abroad. Signs and significance of coefficient estimates are identical to those in Table 2 and remain strongly consistent with positive sorting. The magnitudes of coefficient estimates are now larger, suggesting that positive

[36] Appendix Table C.4 reports the complete regression results.

[37] The data include 16 individual countries (Australia, Brazil, Canada, China, Colombia, England, France, Germany, Hong Kong, Italy, Korea, Mexico, Russia, Poland, Romania, and the US), and six aggregate regions (Central America, Eastern Europe, the Middle East, South America, Southeast Asia, other Western Europe).

[38] See https://www.canada.ca/en/immigration-refugees-citizenship/services/immigrate-canada.

sorting is stronger. Comparing 25^{th} and 75^{th} percentile countries for PISA math scores, the higher scoring country would now have a 106.1% higher share (1.2 of a standard deviation) of its workers in management and finance (the 75^{th} percentile occupation for cognitive task intensity); comparing 25^{th} and 75^{th} percentile countries for linguistic similarity to the US, the more similar country would now have a 15.2% higher share (0.2 of a standard deviation) of its workers in executive management (the 75^{th} percentile occupation for communication intensity). Turning to conditions (A′), (B′), and (C′), as shown in Table 3, we reject that $\beta_{cog}\gamma_{com} - \beta_{com}\gamma_{cog} = 0$ (or that the product of sorting within natural skill dimensions equals the product of sorting across natural skill dimensions) at a 7% significance level; we again reject the joint hypothesis that (A′), (B′) and (C′) at any significance level.

In panel II of Table 6, we examine immigrants who arrived in the US at age 17 or younger. These individuals likely migrated at the behest of their parents or another family member. Since they arrived in the US before age 18, they would have completed at least part of their K-12 education in the US. For this group, the coefficient on the PISA math score and cognitive task intensity interaction is small and highly imprecisely estimated ($\widehat{\beta}_{cog}$ = 0.53, t-value = 0.20). When foreign-born workers undertake K-12 education outside of their country of origin, there is no longer a connection between cognitive skill (as measured by international assessments) and worker sorting across occupations in the US. The results are consistent with the quality of educational institutions in origin countries (rather than origin-country cultural values) playing a determinative role in occupational choice by their workers abroad.

Panel II of Table 6 reveals a positive effect ($\widehat{\gamma}_{com}$ =1.98, t-value = 2.11) for the interaction of linguistic proximity to the US and communication task intensity. Similar to immigrants who arrived as adults, linguistic proximity to the US affects occupational sorting for immigrants who arrived as children. Those from countries with weaker facility in English are less likely to select more communication-intensive jobs. Bleakley and Chin (2004) find that among US immigrants from non-English-speaking countries, those arriving before age 14 tend to have stronger fluency in English than those arriving at age 14 or later. They attribute this result to changes in brain function at puberty, which occur around age 14 and which diminish the ability to learn new languages. Motivated by their findings, in Appendix Table C.5 we restrict the sample of immigrants to those who arrived in the US at age 13 or younger. We continue to find a small and insignificant interaction between PISA scores and cognitive task intensity ($\widehat{\beta}_{cog}$ = 0.57, t-value = 0.19). The interaction between linguistic proximity to the US and communication intensity becomes smaller and statistically insignificant ($\widehat{\gamma}_{com}$ = 1.65, t-value = 1.56). These results suggest, intuitively, that immigrants arriving in the US as young children are less affected by English-language proficiency in their birth countries.

Table 6: OLS results for $\log \Pi^s_{d,o}$ using DOT task intensities, by immigrant age of arrival and native-born ancestry

	(1) Average math	(2) 75ptl math	(3) 90ptl math
	Panel I: Immigrants arriving at age 18+		
$Cog_s \times cog_o$	10.62	11.03	12.72
	(3.063)	(2.564)	(2.393)
$Cog_s \times com_o$	2.112	1.461	2.346
	(2.364)	(2.722)	(2.469)
$Ling_s \times cog_o$	0.481	0.530	0.470
	(0.510)	(0.566)	(0.544)
$Ling_s \times com_o$	1.062	1.024	1.026
	(0.454)	(0.468)	(0.453)
Observations	1691	1496	1633
Adjusted R^2	0.355	0.358	0.358
Number of countries	69	61	67
	Panel II: Immigrants arriving before age 18		
$Cog_s \times cog_o$	0.526	0.364	1.677
	(2.597)	(2.386)	(2.388)
$Cog_s \times com_o$	1.375	2.717	2.429
	(2.710)	(2.808)	(2.753)
$Ling_s \times cog_o$	-0.480	-0.584	-0.560
	(0.657)	(0.642)	(0.654)
$Ling_s \times com_o$	1.981	1.962	1.872
	(0.936)	(0.971)	(0.958)
Observations	1350	1193	1294
Adjusted R^2	0.259	0.257	0.266
Number of countries	68	60	66
	Panel III: Native born by region of ancestry		
$Cog_s \times cog_o$	2.945	3.693	4.383
	(1.993)	(1.513)	(1.358)
$Cog_s \times com_o$	1.047	2.449	1.914
	(2.782)	(3.121)	(2.747)
$Ling_s \times cog_o$	-0.0289	-0.0584	-0.0810
	(0.371)	(0.298)	(0.315)
$Ling_s \times com_o$	-0.777	-0.784	-0.878
	(0.225)	(0.215)	(0.255)
Observations	795	714	743
Adjusted R^2	0.252	0.284	0.252
Number of countries	28	25	26

Notes: Standard errors are clustered by origin country and reported in parentheses. The full set of regressors is the same as for the regressions reported in Table 2.

In panel III of Table 6, we examine US-born individuals by their region of ancestry. Comparing results with panel I, the interaction between PISA math scores in the ancestral region and cognitive task intensity is approximately one third the magnitude; it is imprecisely estimated for average PISA scores and precisely estimated for 75^{th} and 90^{th} percentile scores. The interaction between linguistic proximity to the US and communication task intensity is now *negative*, in contrast to the panel I results. Summarizing results across panels in Table 6, evidence of positive sorting according to cognitive and linguistic skills of the origin country is much stronger for immigrants who arrive as adults than for immigrants who arrived as children or whose parents or older generations where those who migrated to the US.

In Appendix Table C.6, we replicate the results in Table 6, now using O*NET task intensities in place of DOT task intensities. For immigrants arriving at age 18 or older, shown in panel I, we again find positive and statistically significant interactions between PISA test scores and cognitive task intensity and between linguistic proximity and communication task intensity. Similar to the above, neither result holds for immigrants arriving before age 18 (panel II) or native-born individuals identified by their region of ancestry (panel III).

4.4 Extended Results

4.4.1 Temporary Immigration

In a further exercise, we examine the importance of time spent in the US for immigrant sorting across occupations. One reason tenure in the US may matter has to do with rules governing the temporary immigration of skilled workers. In the years up to and including our sample period, the US authorized 65,000 to 85,000 H-1B visas per year to foreign-born workers in specialty occupations. The visas allow a recipient to remain in the US for three years with one three-year extension.[39] In practice, the majority of these visas go to workers in computing and engineering occupations who are employed by major technology companies (Bound et al., 2015, 2017). The role of H-1B visas in tracking foreign-born workers into technology-related jobs could imply that visa allocation and work rules are distorting occupational sorting and thereby diminishing the effect of comparative advantage.

Given that the maximum duration of stay for workers on H-1B visas is six years, we rerun our analysis for immigrants who have been in the US for 7 years or more. The results, in Appendix Table C.7 for all immigrants and Appendix Table C.8 for immigrants who arrived in the US at age 18 or older, are very similar to those already reported. We interpret these findings to mean that temporary work visas are unlikely to be a significant confound.

[39]This implies that that the number of individuals in the US on H-1B visas an any moment in time should not exceed 510,000. Because the US government does not track the location of H-1B visa holders, there is no official count of the stock of US workers who hold H-1B visas.

4.4.2 Occupational Downgrading

A second reason that tenure in the US may be related to job choice is the phenomenon of occupational downgrading among recent arrivals in a new labor market (Dustmann et al., 2013). If immigrants' origin-country training or work experience do not match the specific needs of US employers or meet US occupational licensing requirements, it may take new immigrants time to work their way into a job that is commensurate with their skills. In the case of Russian Jews migrating to Israel following the dissolution of the Soviet Union—a population in which the strong majority of individuals had some post-secondary education—Eckstein and Weiss (2004) report substantial occupational downgrading in the years immediately after arrival followed by steady occupational upgrading over the ensuing decade. The results in Appendix Tables C.7 and C.8, for the sample of immigrants that have at least six years of tenure in the US, suggest that whatever temporary occupational downgrading takes place is unlikely to negate the role of comparative advantage in sorting across jobs.

Nevertheless, the question remains whether college-educated immigrants are subject to occupational downgrading of some kind. Evaluating downgrading is complicated by the fact that the ACS does not report an immigrant's occupation prior to coming to the US. However, it does report the first and second university degree fields for individuals who completed four years of college or an advanced degree. This allows us to evaluate the matching of degree holders to occupations for the subset of degree fields that track individuals into specific lines of work. Among immigrants in our sample, the most common degree fields in descending order of importance are engineering, business, computer and information sciences, social sciences, physical sciences, and biology and life sciences, which together account for 67.3% of primary degrees held by immigrants in our sample. Whereas engineering and computer information systems map to specific jobs, business and social sciences do not. Biological and physical sciences would appear to be somewhere in between.

We evaluate occupational matching for immigrants whose primary or secondary degree fields belong to the narrower category of computer-engineering (which includes degrees in computer information systems, computer science, and engineering),[40] and the broader category of STEM (which includes computer-engineering plus mathematics and biological and physical sciences).[41] We examine whether origin-country PISA scores or linguistic similarity to the US affect the likelihood that individuals who majored in these fields take degree-appropriate jobs. We match computer-engineering degrees with jobs in three occupations:

[40]These titles correspond to ACS general degree fields, which contain the following detailed degree fields: computer programming, computer science, information science, computer information management, and computer networking; all engineering fields; and mathematics and computer science.

[41]The additional detailed degree fields in the STEM category include: mathematics, applied mathematics, and statistics and decision science; biology, biochemical sciences, botany, molecular biology, ecology, genetics, zoology, neuroscience, and cognitive science; and astronomy, astrophysics, atmospheric sciences, chemistry, geology, geosciences, oceanography, physics, and materials science.

computer software developers, computer system analysts, and engineers; we match STEM degrees to these three occupations plus mathematicians and scientists. Evaluating matching has meaning only relative to some base occupation, which we select to be less-skilled occupations for which no college degree would seem to be required: construction, transportation, machine operation, mechanics and repairers, low-skill services, and mid-skill services. In Table 4, we see that employment in these occupations is relatively low in countries with high PISA scores and greater linguistic proximity to the US. The regression specification is,

$$\log \Pi^s_{d,o} - \log \Pi^s_{d,o'} = \alpha + \beta \text{Cog}_s + \gamma \text{Ling}_s + \phi \text{Dist}_s + \varepsilon_s, \tag{9}$$

where $\log \Pi^s_{d,o} - \log \Pi^s_{d,o'}$ is the log employment share in computer-engineering or STEM occupations, indexed by o, relative to the log employment share for the aggregate of low-skill occupations, indexed by o', for the sample of foreign-born workers whose primary or secondary degree corresponds to occupation o.[42] We report results for all immigrants, those who arrived in the US at age 18 or older, and older arrivals with at least 7 years in the US.

Appendix Table C.9 reports estimation results for (9). We begin with all immigrants in panel I. In column (1), we see that computer-engineering degree holders from higher PISA score countries are more likely to be employed in computer-engineering occupations (than in low-skilled occupations) when compared to computer-engineering degree holders from lower PISA score countries ($\widehat{\beta}$ = 8.25, t-value = 1.95). Similar results obtain for the broader STEM occupational/degree category in column (2). These findings are consistent with two interpretations. One is relative occupational downgrading (individuals from high PISA score countries are more likely to work in occupations that utilize their computer-engineering or STEM training) and another is occupational comparative advantage (individuals from high PISA score countries are more likely to work in cognitive-task-intensive occupations).

To distinguish between these explanations, in column (3) we compare employment in high-skill, non-STEM occupations to employment low-skill occupations, for individuals whose primary or secondary university degree is in a STEM field. If occupational downgrading accounts for the results in the first two columns of Appendix Table C.9, then we should find much smaller impacts of PISA scores on selection into high-skill, non-STEM occupations (since STEM degrees are not essential for these jobs). Coefficient estimates are very similar in columns (1), (2), and (3), which indicates that individuals from high PISA score countries (with university degrees in STEM fields) are relatively likely to select into high-skill jobs, be they in STEM or non-STEM occupations. A similar pattern holds for immigrants arriving in the US at age 18 or older, in panel II, and for these older immigrants with at least seven

[42]Because we are comparing single pairs of occupations across workers grouped by origin country, we cannot include origin-country or occupation fixed effects in the regression. It is further the case that differences in task intensities between the two occupations are absorbed into the regression coefficients.

years in the US, in panel III. We interpret these results as confirming our earlier findings on comparative advantage and as unsupportive of relative occupational downgrading.

5 Counterfactual Exercises

Although understanding the role of origin-country occupational comparative advantage in immigrant sorting across jobs is worthwhile in its own right, the estimation results are also of use for evaluating the labor-market consequences of changing the skill bias of US immigration policy. We perform counterfactual exercises in which we adjust US visa policies to place greater weight on observable dimensions of skill.

5.1 A General Equilibrium Model

We begin by incorporating the labor supply side of the economy that we have examined thus far into a general equilibrium model. For simplicity, we abstract away from international trade in goods or capital flows between countries. We assume that the US produces a final good by aggregating inputs supplied by each occupation using a CES technology. We incorporate non-college educated workers into the model, such that s now indexes the origin country and the education level (less than college, college degree or higher). In one version of the model, we assume that workers from different origin-country and education groups s are perfect substitutes within each occupation. The CES production function is

$$Q_d = \left[\sum_o A_{d,o} L_{d,o}^{\frac{\rho-1}{\rho}} \right]^{\frac{\rho}{\rho-1}},$$

where $A_{d,o}$ is the exogenous factor productivity in occupation o, ρ denotes the elasticity of substitution across occupations, and $L_{d,o}$ is total efficiency units of labor employed in occupation o. In a second version of the model, presented in Appendix D.2, we assume that native and foreign-born workers are imperfect substitutes within each occupation based on a nested CES production function (where the foreign born of different education groups are perfectly substitutable among themselves). Following Burstein et al. (2020), we set $\rho = 1.6$; and we set the Fréchet dispersion parameter $\theta = 4$.[43]

In order to make US immigration policy responsive to skill, we model migration costs as a decreasing function of PISA exam scores. In so doing, we make the US more skill biased in its visa allocations. Modelling immigration policy in this way requires that we characterize the equilibrium labor allocation within each country at more disaggregate level. The fraction

[43] In the the literature, the calibrated or estimated values of θ range between 1.5 and 5.3. For example, Lagakos and Waugh (2013) obtain a value of 5.3 for agriculture, and 2.7 for the non-agricultural sector; Hsieh et al. (2013) obtain values between 3.4 and 4.6; and Burstein et al. (2015) obtain a value equal to 1.8. Our value of θ is somewhat at the high end for the literature. In this way, we ensure that selection bias in immigration for an origin country (captured by the term $\Pi_d^s(p)^{-\frac{1}{\theta}}$, as derived below) is relatively insensitive to changes in the magnitude of immigration. This allows us to focus on labor-market impacts of changing the composition of immigrants across origin countries (rather than changes in immigrant composition from a given origin).

of workers from origin country s with PISA score p who live in country d is given by,

$$\Pi_d^s(p) = \frac{\sum_o \left[T_{d,o}^s(p)\tau_{d,o}^s(p)w_{d,o}\right]^\theta}{\sum_{d',o'} \left[T_{d',o'}^s(p)\tau_{d',o'}^s(p)w_{d',o'}\right]^\theta}. \tag{10}$$

Aggregating across individuals, the fraction of group s workers who live in country d is,

$$\Pi_d^s = \int \frac{\sum_o \left[T_{d,o}^s(p)\tau_{d,o}^s(p)w_{d,o}\right]^\theta}{\sum_{d',o'} \left[T_{d',o'}^s(p)\tau_{d',o'}^s(p)w_{d',o'}\right]^\theta} dF_p^s \tag{11}$$

where F_p^s is the distribution of PISA scores among individuals in group s. Although we do not observe the full distribution of scores for each country, for 2015 we do have the distribution of scores across seven proficiency levels (from the National Center for Education Statistics).[44] We discretize the distribution of scores to match the available data. After merging PISA scores for 2015 into the ACS, we have data on 58 countries. Appendix Table C.10 reports summary statistics on PISA math scores by proficiency level.

Figure 5 plots the fraction of students who achieved the two highest proficiency levels (5 or 6) in the PISA math exam in 2015 against the fraction that achieved the third highest level (4) in that year. Summing the values on the vertical and horizontal axes for each country reveals that in most countries less than 30 percent of students achieve scores in the top three performance categories; in Appendix Table C.10, countries on average have 23.5% of their students earn scores in one of the top three levels. The fraction of high-achieving students is relatively high in East Asian countries—including Singapore, China, Japan, Korea, and Taiwan—and relatively low in Latin American countries—including Mexico, Brazil, Peru, and Colombia. For instance, 60.0% of Singaporean students and 47.4% of Chinese students have proficiency levels of 4 or above, whereas the comparable figures are 3.2% for Mexico and 2.4% for Colombia. Making US immigration policy responsive to PISA scores would have the effect of reallocating visas in favor Asian countries.

[44]Level 0 is a score less than or equal to 334.9; level 1 is a score greater than 334.9 and less than or equal to 409.5; level 2 is a score greater than 409.5 and less than or equal to 484.1; level 3 is a score greater than 484.1 and less than or equal to 558.7; level 4 is a score greater than 558.7 and less than or equal to 633.3; level 5 is a score greater than 633.3 and less than or equal to 707.9; and level 6 is a score greater than 707.9.

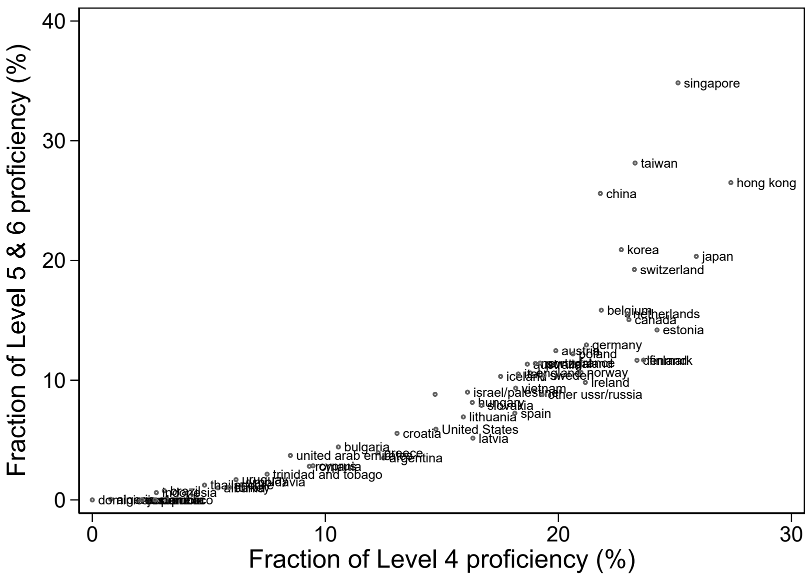


Figure 5: Student performance on PISA math exams, 2015

One shortcoming of PISA scores is that of the 115 origin countries represented among foreign-born, prime-age, college-educated males working in the US, we only have exam scores for 58 of these. Of the excluded countries, India is the most important. Whereas India accounts for 20% of foreign-born, college-educated men in our sample, the other 56 countries absent in the PISA data account for just 28%. To address the omission of India, we use a synthetic control approach (Abadie and Gardeazabal, 2003) to impute a PISA score distribution for the country. Details are in Appendix D.1.

5.2 Simulating Counterfactuals

Denote $\widehat{\tau}_d(p)$ as the assumed proportional change in migration costs for workers with PISA exam score p. Among the college-educated, we increase $\widehat{\tau}_d(p)$ for workers whose value of p belongs to proficiency levels 4 to 6 and reduce $\widehat{\tau}_d(p)$ for workers whose value of p belongs to proficiency levels 0 to 3. From Appendix Table C.10, this policy is roughly equivalent to increasing visa allocations for individuals with PISA scores above the 75^{th} percentile and reducing visa allocations for everyone else. We assume that migration costs change only among college-educated workers from the 58 countries (plus India) for which we have PISA score distributions; we keep migration costs unchanged for other countries. We also hold constant migration costs among non-college-educated workers. We select magnitudes for $\widehat{\tau}_d(p)$ that yield the result of halving the stock of college-educated workers with a PISA proficiency level of 0 to 3, while increasing the stock of those with proficiency level of 4 to 6 in a manner that leaves the total stock of US college-educated immigrants unchanged.

The effect of this policy change is to reallocate US immigration visas across origin countries. Because fundamental comparative advantage, as summarized by the Fréchet scale parameter $T^s_{d,o'}$ differs across countries, changes in visa allocations will in turn reallocate labor across US occupations. Employment will tend to expand in occupations in which countries awarded more visas have a comparative advantage and contract in occupations for which the opposite is true. Within origin countries, the change in visa policy would potentially reallocate visas from individuals with low PISA scores to individuals with high PISA scores. Given that we do not observe exam scores for individuals in the ACS, the data make it difficult for us to account for changes in the composition of immigrants within origin countries. These data limitations necessitate additional assumptions, which we describe below.

To recover the change in equilibrium outcomes under the new visa policy that we impose, we apply the Exact Hat Algebra of Dekle et al. (2008) to solve the model for the US economy. To simplify the analysis, we assume that changes in US migration costs do not affect equilibrium wage values in immigrant-sending countries.[45] The proportional change in the migration rate for those with PISA score p is,

$$\widehat{\Pi}^s_d(p) = \frac{\widehat{\tau}^\theta_d(p)\Big[\sum_o \widehat{w}^\theta_{d,o}\Pi^s_{o|d}(p)\Big]}{\sum_{d'} \widehat{\tau}^\theta_{d'}(p)\Big[\sum_{o'} \widehat{w}^\theta_{d',o'}\Pi^s_{o|d}(p)\Big]\Pi^s_{d'}(p)} \approx \frac{\widehat{\tau}^\theta_d(p)\Big[\sum_o \widehat{w}^\theta_{d,o}\Pi^s_{o|d}(p)\Big]}{\widehat{\tau}^\theta_d(p)\Big[\sum_o \widehat{w}^\theta_{d,o}\Pi^s_{o|d}(p)\Big]\Pi^s_d(p) + 1 - \Pi^s_d(p)} \tag{12}$$

where $\Pi^s_{o|d}(p)$ is the fraction of immigrants from group s with score p working in US occupation o; and $\Pi^s_d(p)$ is the migration rate to the US among workers in s with score p. The change in the aggregate migration rate for workers from origin group s is then,

$$\widehat{\Pi}^s_d = \int \Pi^s_d(p)\widehat{\Pi}^s_d(p)\, dF^s_p \Big/ \Pi^s_d.$$

In this setting, countries that have a fatter tail in their PISA score distribution will see larger changes in labor flows. The proportional change in the fraction of workers from origin country s who work in occupation o with PISA score p is,

$$\widehat{\Pi}^s_{o|d}(p) = \frac{\widehat{w}^\theta_{d,o}}{\sum_{o'} \widehat{w}^\theta_{d,o'}\Pi^s_{o|d}(p)}. \tag{13}$$

[45] Allowing for wage adjustments in immigrant-sending countries would have little impact on our results. In our counterfactual, $\widehat{\tau}_d(p)$ falls by 16.5% for proficiency levels 0-3, while increasing 25.1% for proficiency levels 4-6. The wage impacts in immigrant-sending countries resulting from these US immigration policy changes would be at most 3% (Liu, 2020). Since migration costs and wage units enter the migration allocation formula symmetrically, origin-country wage adjustments would have very small impacts on our results.

Changes in aggregate occupational labor demand and labor supply are

$$\widehat{L}_{d,o}^{\text{demand}} = \widehat{w}_{d,o}^{-\rho}\widehat{Y}_d\widehat{A}_{d,o}^{\rho},\text{[46]} \tag{14}$$

$$\widehat{L}_{d,o}^{\text{supply}} = \frac{1}{\widehat{w}_{d,o}}\Big[\sum_s \int W_{d,o}^s(p)\cdot L_d^s(p)\cdot \Pi_{o|d}^s(p)\cdot \widehat{W}_{d,o}^s(p)\cdot \widehat{\Pi}_d^s(p)\cdot \widehat{\Pi}_{o|d}^s(p)\, dF_p^s\Big] \Big/ \sum_s W_{d,o}^s \cdot L_d^s \cdot \Pi_{o|d}^s, \tag{15}$$

where

$$\widehat{W}_{d,o}^s(p) = \widehat{w}_{d,o}\Big(\widehat{\Pi}_{o|d}^s(p)\widehat{\Pi}_d^s(p)\Big)^{-\frac{1}{\theta}}, \tag{16}$$

is the change in average occupational wages for workers from origin group s with PISA score p. We solve $\widehat{w}_{d,o}$ such that $\widehat{L}_{d,o}^{\text{demand}} = \widehat{L}_{d,o}^{\text{supply}}$.

Given that we do not observe PISA scores at the individual level, we are obligated to make additional assumptions in order to solve equations (13) to (16). First, because we do not observe occupation employment shares by PISA score, we assign the same share $\Pi_{o|d}^s(p)$ to all workers in group s. By (13) and (16), the proportional change in occupation employment is then the same across all p for workers from s, and the proportional change in averages wages for workers from s does not depend on p. This assumption means that the policy change reallocates visas from high-PISA-score to low-PISA-score countries, without changing the skill distribution of immigrants from a country. Second, and relatedly, because we do not observe occupational wages by PISA score, we assume that the equilibrium occupational wage is constant across p for immigrants from s. All college-educated immigrants from s thus have the same efficiency units. Third, we assume the migration probability $\Pi_d^s(p)$ does not vary across workers from s, which neutralizes selection on unobservables in terms of the PISA score in determining the migration propensity. By (12), the proportional changes in migration rates is then the same across all p for immigrants from s.

The net effect of these assumptions is to neutralize variation in policy impacts across workers with a given education level from a given origin country and thereby concentrate impacts on the variation in outcomes across origin countries. Similar to our empirical analysis, we address positive sorting across but not within origin countries for US immigration. Employment and wage outcomes by occupation are purely a response to policy-induced changes in national origin of US immigrants. If there is positive sorting across occupations among workers from each origin country (e.g., if workers with high PISA scores are more likely to work in occupations intensive in cognitive tasks), then our counterfactual exercises would underestimate policy impacts on occupational wages and employment.

[46] Where the change in aggregate income is,

$$\widehat{Y}_d = \frac{\sum_s \sum_o W_{d,o}^s \cdot L_d^s \cdot \Pi_{o|d}^s \cdot \widehat{W}_{d,o}^s \cdot \widehat{\Pi}_d^s \cdot \widehat{\Pi}_{o|d}^s}{\sum_s \sum_o W_{d,o}^s \cdot L_d^s \cdot \Pi_{o|d}^s}.$$

5.3 Quantitative Results

The results for our counterfactual exercises appear in Tables 7 and 8. When allowing visas for India to adjust, as shown in the second row of the Table 7, the change in US visa policy increases the overall employment of college-educated immigrants from China (gain of 2.3 percentage points), India (gain of 2.4 percentage points), and elsewhere in East Asia (gain of 2.7 percentage points), while reducing the presence of college-educated immigrants from other regions (loss of 0.6 percentage points), Southeast Asia (loss of 0.9 percentage points), and Latin America (loss of 6.1 percentage points). When holding immigration from India constant, as shown in the third row of Table 7, the gains in immigration shares for China and other East Asia expand, while there is little change for other regions.

Table 7: Share of college-educated US immigrants by origin country

China	EU	Other East Asia	India	Latin America	Other regions	SE Asia
Data in 2010						
7.2%	6.7%	9.9%	16.4%	23.9%	21.6%	14.1%
Counterfactual outcomes with perfect substitution, allowing India to change						
9.5%	6.9%	12.6%	18.8%	17.8%	21.0%	13.3%
Counterfactual outcomes with perfect substitution, holding India constant						
10.1%	7.3%	13.4%	16.4%	17.9%	21.5%	13.5%

Notes: Values are fractions of immigrants from each origin region as as share of all US immigrants, among male, prime-age, college-educated, foreign-born workers.

These changes in the origin countries for immigration naturally change the skill composition of the US labor force and lead to a reallocation of workers across occupations, which we show in Appendix Table C.11 under perfect substitution between foreign and native-born workers and in Appendix Table C.12 under imperfect substitution between the two groups. With perfect substitution, there are gains in employment, measured in total efficiency units of labor, in four occupations: computer software developers, engineers, scientists and mathematicians, and post-secondary teachers. All other occupations see contractions in employment, with the largest declines occurring for primary and secondary teachers, clergy and the arts, low-skill services, health professionals, and other managers. Intuitively, all occupations in which employment expands see reductions on wages per efficiency unit, whereas most occupations in which employment contracts see increases in wages per efficiency unit.[47] With imperfect substitution, we see falling wages per efficiency unit for immigrants in occupations in which immigrant employment expands (computer software development, engineers, scientists and mathematicians, post-secondary teachers), and, because of imperfect substitu-

[47]Some occupations (e.g., business analysts) see declines both in employment and wages per efficiency unit. Although an employment decline for business analysts alone would push up its occupational wage, the equilibrium wage effect depends on employment changes in the occupation relative to employment changes in other occupations. If the relative decline in employment of business analysts is small, wages per efficiency unit for business analysts will tend to fall.

tion, rising employment for native-born workers in these occupations. In the occupations in which foreign or native-born employment expands, wages per efficiency unit decline, with wages tending to increase in occupations in which efficiency units contract.

Table 8: Impacts on U.S. Native Wages

Perfect Substitution Between Immigrants and Natives		**Imperfect Substitution Between Immigrants and Natives**	
(1)	(2)	(3)	(4)
Non-College	College	Non-College	College
0.0119%	0.0012%	-0.0228%	-0.0235%

Notes: All values are in percentage terms for changes for native wages. The reported values are calculated as $(\widehat{W}_d^s - 1) \times 100$, where $\widehat{W}_d^s$ is the group-level wage changes. For US natives, the average group-level wage changes can be expressed as $\sum_o \left(\hat{w}_d^{\theta} \Pi_{o|d}^s\right)^{\frac{1}{\theta}}$.

Results for changes in the average wages of foreign and native-born workers appear in Table 8. Under perfect substitution of native and foreign-born labor, Appendix Table C.11 shows that wages per efficiency unit increase in most occupations. Because native-born workers are relatively evenly distributed across occupations, it is the case that wages rise on average for both non-college and college-educated native-born workers. We also see that wage increases are larger for the non-college-educated than for college-educated natives, which is due to the fact college-educated natives are more likely to work in occupations (computer software development, post-secondary teachers) that experience increases in immigrant employment and decreases in wages per efficiency unit. Under imperfect substitution between the foreign and native-born, immigrant employment declines in most occupations which, because of imperfect substitution, causes native wages to fall in most occupations. Again because native-born workers are relatively evenly distributed across jobs, on average native-born wages decline, both among the the non-college-educated and the college-educated. Given the simplified nature of our model (e.g., no foreign trade or investment), we view these results as illustrative of the sectoral and distributional impacts of changes in immigration policy, rather than precisely indicative of the outcomes that would occur were US immigration policy to become more skill biased.

6 Conclusion

Observers have long noticed that upon arriving in a new country, immigrants often congregate in occupations according to their country of origin. The standard explanation for such occupational clustering is the presence of migration networks. Early arrivals from a given origin country just happen to choose one set of jobs over another, and, because job search is costly for new arrivals and information flows relatively freely within origin-country migrant communities, later cohorts of immigrants tend to follow in the footsteps of the pioneers. No doubt, migration networks have been a powerful force in immigrant job choice in many historical episodes. Yet, such networks provide a less compelling explanation for job search among the highly educated. Individuals choose to become computer programmers not because it seems like the obvious thing to do but because their training and aptitude makes such a difficult career choice feasible. In jobs in which cognitive reasoning and analytical skill are required, the quality of educational institutions in a country likely affect the career opportunities that individuals from the country have when choosing to migrate abroad.

We present evidence consistent with national comparative advantage in immigrant job choice and with positive sorting in multidimensional skill across jobs. US immigrants from countries whose students score more highly on international assessments—as indicative of the quality of educational institutions in the origin country—specialize more strongly in jobs that are more intensive in cognitive skill. Similarly, US immigrants from countries that are more linguistically similar to the US are more concentrated in jobs that are intensive in interpersonal communication. One implication of these patterns is that changes in the origin-country-bias of US immigration policy—whether implicit or explicit—would change the relative supply of labor across occupations and therefore US employment in these occupations. In effect, the US can choose its occupational comparative advantage by changing how it allocates immigration visas across countries of origin of migrants (at least among those who arrive as adults). Favoring countries that achieve higher test scores in awarding visas is one path to this outcome. Introducing an explicit point system for immigrant admissions, as in Canada, is another path. As the US ponders comprehensive reform to its immigration policies, its comparative advantage across occupations is thus in play.

References

Abadie, A. and Gardeazabal, J. (2003), 'The economic costs of conflict: A case study of the basque country', *American economic review* **93**(1), 113–132.

Acemoglu, D. and Autor, D. (2011), Skills, tasks and technologies: Implications for employment and earnings, *in* 'Handbook of labor economics', Vol. 4, Elsevier, pp. 1043–1171.

Akyol, Ş. P., Krishna, K. and Wang, J. (2018), Taking pisa seriously: How accurate are low stakes exams?, Technical report, National Bureau of Economic Research.

Antecol, H., Cobb-Clark, D. A. and Trejo, S. J. (2003), 'Immigration policy and the skills of immigrants to australia, canada, and the united states', *Journal of Human Resources* **38**(1), 192–218.

Autor, D. and Dorn, D. (2013), 'The growth of low-skill service jobs and the polarization of the us labor market', *American Economic Review* **103**(5), 1553–97.

Autor, D. H. and Handel, M. J. (2013), 'Putting tasks to the test: Human capital, job tasks, and wages', *Journal of labor Economics* **31**(S1), S59–S96.

Autor, D. H., Levy, F. and Murnane, R. J. (2003), 'The skill content of recent technological change: An empirical exploration', *The Quarterly Journal of Economics* pp. 1279–1333.

Beaudry, P., Green, D. A. and Sand, B. M. (2016), 'The great reversal in the demand for skill and cognitive tasks', *Journal of Labor Economics* **34**(S1), S199–S247.

Bharadwaj, P., De Giorgi, G., Hansen, D. and Neilson, C. (2012), The gender gap in mathematics: Evidence from low-and middle-income countries, Technical report, National Bureau of Economic Research.

Bleakley, H. and Chin, A. (2004), 'Language skills and earnings: Evidence from childhood immigrants', *The Review of Economics and Statistics* **86**(2), 481–496.

Borjas, G. J. (1987), 'Self-selection and the earnings of immigrants', *The American Economic Review* pp. 531–553.

Borjas, G. J. (1993), Immigration policy, national origin, and immigrant skills: A comparison of canada and the united states, *in* 'Small Differences That Matter: Labor Markets and Income Maintenance in Canada and the United States', University of Chicago Press, pp. 21–44.

Borjas, G. J. and Doran, K. B. (2012), 'The collapse of the soviet union and the productivity of american mathematicians', *The Quarterly Journal of Economics* **127**(3), 1143–1203.

Bound, J., Braga, B., Golden, J. M. and Khanna, G. (2015), 'Recruitment of foreigners in the market for computer scientists in the united states', *Journal of labor economics* **33**(S1), S187–S223.

Bound, J., Khanna, G. and Morales, N. (2017), Understanding the economic impact of the h-1b program on the us, *in* 'High-Skilled Migration to the United States and its Economic Consequences', University of Chicago Press.

Burstein, A., Hanson, G., Tian, L. and Vogel, J. (2020), 'Tradability and the labor-market impact of immigration: Theory and evidence from the united states', *Econometrica* **88**(3), 1071–1112.

Burstein, A., Morales, E. and Vogel, J. (2015), 'Accounting for changes in between-group inequality', *National Bureau of Economic Research* .

Chiquiar, D. and Hanson, G. H. (2005), 'International migration, self-selection, and the distribution of wages: Evidence from mexico and the united states', *Journal of political economy* **113**(2), 239–281.

Chiswick, B. R. and Taengnoi, S. (2007), 'Occupational choice of high skilled immigrants in the united states', *International Migration* **45**(5), 3–34.

Chor, D. (2010), 'Unpacking sources of comparative advantage: A quantitative approach', *Journal of International Economics* **82**(2), 152–167.

Cortés, P. and Pan, J. (2014), 'Foreign nurse importation and the supply of native nurses', *Journal of Health Economics* **37**, 164–180.

Costinot, A. and Vogel, J. (2010), 'Matching and inequality in the world economy', *Journal of Political Economy* **118**(4), 747–786.

Costinot, A. and Vogel, J. (2015), 'Beyond ricardo: Assignment models in international trade', *Annual Review of Economics* **7**(1), 31–62.

Dekle, R., Eaton, J. and Kortum, S. (2008), 'Global rebalancing with gravity: Measuring the burden of adjustment', *IMF Staff Papers* **55**(3), 511–540.

Deming, D. J. (2017), 'The growing importance of social skills in the labor market', *The Quarterly Journal of Economics* **132**(4), 1593–1640.

Deming, D. J. and Noray, K. (2020), 'Earnings dynamics, changing job skills, and stem careers', *The Quarterly Journal of Economics* **135**(4), 1965–2005.

Dustmann, C. and Fabbri, F. (2003), 'Language proficiency and labour market performance of immigrants in the uk', *The Economic Journal* **113**(489), 695–717.

Dustmann, C., Frattini, T. and Preston, I. P. (2013), 'The effect of immigration along the distribution of wages', *Review of Economic Studies* **80**(1), 145–173.

Eckstein, Z. and Weiss, Y. (2004), 'On the wage growth of immigrants: Israel, 1990–2000', *Journal of the European Economic Association* **2**(4), 665–695.

Federman, M. N., Harrington, D. E. and Krynski, K. J. (2006), 'Vietnamese manicurists: are immigrants displacing natives or finding new nails to polish?', *ILR Review* **59**(2), 302–318.

Fernandez, R. and Fogli, A. (2009), 'Culture: An empirical investigation of beliefs, work, and fertility', *American economic journal: Macroeconomics* **1**(1), 146–77.

Figlio, D., Giuliano, P., Özek, U. and Sapienza, P. (2019), 'Long-term orientation and educational performance', *American Economic Journal: Economic Policy* **11**(4), 272–309.

Figlio, D. and Özek, U. (2020), 'Cross-generational differences in educational outcomes in the second great wave of immigration', *Education Finance and Policy* pp. 1–68.

Fryer, R. G. and Levitt, S. D. (2010), 'An empirical analysis of the gender gap in mathematics', *American Economic Journal: Applied Economics* **2**(2), 210–240.

Gelatt, J. (2020), 'Do employer-sponsored immigrants fare better in labor markets than family-sponsored immigrants?', *RSF: The Russell Sage Foundation Journal of the Social Sciences* **6**(3), 70–93.

Goos, M. and Manning, A. (2007), 'Lousy and lovely jobs: The rising polarization of work in britain', *The review of economics and statistics* **89**(1), 118–133.

Goos, M., Manning, A. and Salomons, A. (2014), 'Explaining job polarization: Routine-biased technological change and offshoring', *American economic review* **104**(8), 2509–26.

Grogger, J. and Hanson, G. H. (2011), 'Income maximization and the selection and sorting of international migrants', *Journal of Development Economics* **95**(1), 42–57.

Guiso, L., Monte, F., Sapienza, P. and Zingales, L. (2008), 'Diversity. culture, gender, and math.', *Science (New York, NY)* **320**(5880), 1164–1165.

Han, S. and Kleiner, M. M. (2016), Analyzing the influence of occupational licensing duration on labor market outcomes, Technical report, National Bureau of Economic Research.

Hanson, G. H. (2021), Immigration and regional specialization in ai, Working Paper 28671, National Bureau of Economic Research.

Hanson, G. H. and Liu, C. (2017), High-skilled immigration and the comparative advantage of foreign-born workers across us occupations, *in* 'Talent Flows in the Global Economy', University of Chicago Press.

Hanushek, E. A. and Kimko, D. D. (2000), 'Schooling, labor-force quality, and the growth of nations', *American economic review* pp. 1184–1208.

Hanushek, E. A., Kinne, L., Lergetporer, P. and Woessmann, L. (2020), 'Culture and student achievement: The intertwined roles of patience and risk-taking', *NBER Working Paper* (w27484).

Hanushek, E. A. and Woessmann, L. (2011), The economics of international differences in educational achievement, *in* 'Handbook of the Economics of Education', Vol. 3, pp. 89–200.

Head, K. and Mayer, T. (2013), 'What separates us? sources of resistance to globalization', *Canadian Journal of Economics/Revue canadienne d'économique* **46**(4), 1196–1231.

Head, K. and Ries, J. (2001), 'Increasing returns versus national product differentiation as an explanation for the pattern of us-canada trade', *American Economic Review* **91**(4), 858–876.

Hershbein, B. and Kahn, L. B. (2018), 'Do recessions accelerate routine-biased technological change? evidence from vacancy postings', *American Economic Review* **108**(7), 1737–72.

Hsieh, C.-T., Hurst, E., Jones, C. I. and Klenow, P. J. (2013), 'The allocation of talent and us economic growth', *NBER Working Paper* .

Hsieh, C.-T., Hurst, E., Jones, C. I. and Klenow, P. J. (2019), 'The allocation of talent and us economic growth', *Econometrica* **87**(5), 1439–1474.

Hunt, J. (2011), 'Which immigrants are most innovative and entrepreneurial? distinctions by entry visa', *Journal of Labor Economics* **29**(3), 417–457.

Hunt, J. (2015), 'Are immigrants the most skilled us computer and engineering workers?', *Journal of Labor Economics* **33**(S1), S39–S77.

Hunt, J. and Gauthier-Loiselle, M. (2010), 'How much does immigration boost innovation?', *American Economic Journal: Macroeconomics* **2**(2), 31–56.

Isphording, I. E. and Otten, S. (2014), 'Linguistic barriers in the destination language acquisition of immigrants', *Journal of Economic Behavior & Organization* **105**, 30–50.

Jackson, C. K. and Schneider, H. S. (2011), 'Do social connections reduce moral hazard? evidence from the new york city taxi industry', *American Economic Journal: Applied Economics* **3**(3), 244–67.

Lagakos, D. and Waugh, M. E. (2013), 'Selection, agriculture, and cross-country productivity differences', *American Economic Review* **103**(2), 948–80.

Lavy, V. (2015), 'Do differences in schools' instruction time explain international achievement gaps? evidence from developed and developing countries', *The Economic Journal* **125**(588), F397–F424.

Lazear, E. P. (2021), 'Why are some immigrant groups more successful than others?', *Journal of Labor Economics* **39**(1), 115–133.

Levchenko, A. A. (2007), 'Institutional quality and international trade', *The Review of Economic Studies* **74**(3), 791–819.

Lindenlaub, I. (2017), 'Sorting multidimensional types: Theory and application', *The Review of Economic Studies* **84**(2), 718–789.

Liu, C. (2020), 'Modes of entry, correlated productivity, and the global impacts of us immigration reform', *Correlated Productivity, and the Global Impacts of US Immigration Reform (August 31, 2020)* .

Llull, J. (2018), 'Immigration, wages, and education: A labour market equilibrium structural model', *The Review of Economic Studies* **85**(3), 1852–1896.

McManus, W. S. (1990), 'Labor market effects of language enclaves: Hispanic men in the united states', *Journal of Human Resources* pp. 228–252.

Melitz, J. and Toubal, F. (2014), 'Native language, spoken language, translation and trade', *Journal of International Economics* **93**(2), 351–363.

Mogstad, M., Romano, J. P., Shaikh, A. and Wilhelm, D. (2020), Inference for ranks with applications to mobility across neighborhoods and academic achievement across countries, Technical report, National Bureau of Economic Research.

Munshi, K. (2003), 'Networks in the modern economy: Mexican migrants in the us labor market', *The Quarterly Journal of Economics* **118**(2), 549–599.

Munshi, K. (2020), 'Social networks and migration', *Annual Review of Economics* **12**(1), 503–524.

Novy, D. (2013), 'Gravity redux: measuring international trade costs with panel data', *Economic inquiry* **51**(1), 101–121.

OIS (2020), *Yearbook of Immigration Statistics*, US Department of Homeland Security, Office of Immigration Statistics.

Oreopoulos, P. (2011), 'Why do skilled immigrants struggle in the labor market? a field experiment with thirteen thousand resumes', *American Economic Journal: Economic Policy* **3**(4), 148–71.

Oreopoulosa, P. (2011), 'Why do skilled immigrants struggle in the labor market? a field experiment with thirteen thousand resumes', *American Economic Journal: Economic Policy* **3**(4), 148–171.

Patel, K. and Vella, F. (2013), 'Immigrant networks and their implications for occupational choice and wages', *Review of Economics and Statistics* **95**(4), 1249–1277.

Pellegrina, H. S. and Sotelo, S. (2021), Migration, specialization, and trade: Evidence from brazil's march to the west, Technical report, National Bureau of Economic Research.

Peri, G., Shih, K. and Sparber, C. (2015), 'Stem workers, h-1b visas, and productivity in us cities', *Journal of Labor Economics* **33**(S1), S225–S255.

Peri, G. and Sparber, C. (2009), 'Task specialization, immigration, and wages', *American Economic Journal: Applied Economics* **1**(3), 135–169.

Peri, G. and Sparber, C. (2011), 'Highly educated immigrants and native occupational choice', *Industrial Relations: a journal of economy and society* **50**(3), 385–411.

Romalis, J. (2004), 'Factor proportions and the structure of commodity trade', *American Economic Review* **94**(1), 67–97.

Roy, A. D. (1951), 'Some thoughts on the distribution of earnings', *Oxford economic papers* **3**(2), 135–146.

Sands, G. (2017), 'Are the pisa education results rigged. forbes'.

Schoellman, T. (2012), 'Education quality and development accounting', *The Review of Economic Studies* **79**(1), 388–417.

Shu, X., Fan, P.-L., Li, X. and Marini, M. M. (1996), 'Characterizing occupations with data from thedictionary of occupational titles', *Social Science Research* **25**(2), 149–173.

Silva, J. S. and Tenreyro, S. (2006), 'The log of gravity', *The Review of Economics and statistics* **88**(4), 641–658.

Woessmann, L. (2016), 'The importance of school systems: Evidence from international differences in student achievement', *Journal of Economic Perspectives* **30**(3), 3–32.

Woodruff, C. and Zenteno, R. (2007), 'Migration networks and microenterprises in mexico', *Journal of development economics* **82**(2), 509–528.

Xiang, C. and Yeaple, S. (2018), The production of cognitive and non-cognitive human capital in the global economy, Technical report, National Bureau of Economic Research.

A Appendix

Here, we describe the O*NET variables we use to measure occupational task intensity.

Social task intensity: we measure using the four variables social perceptiveness (being aware of others' reactions and understanding why they react the way they do), coordination (adjusting actions in relation to others' actions), persuasion (persuading others to approach things differently), and negotiation (bringing others together to reconcile differences).

Cognitive task intensity: we measure using the three variables mathematical reasoning ability (ability to understand and organize a problem and then to select a mathematical method or formula to solve the problem), mathematics knowledge (knowledge of numbers, their operations, and interrelationships including arithmetic, algebra, geometry, calculus, and statistics), and mathematics skill (using mathematics to solve problems).

Routine task intensity: we measure using the two variables degree of automation (level of automation of this job), and the importance of repeating tasks (importance of repeating the same physical activities or mental activities over and over, without stopping).

Manual task intensity: we measure using the two variables assisting and caring for others (providing assistance or personal care to others), and service orientation (actively looking for ways to help people).

Table A.1: US employment of prime-age, college-educated, foreign-born men (% of total)

	Aggregate employment share	Employment share by broad occupation					
		Management, accounting, finance	STEM	Health	Education, law, arts, social work	Admin., technical, sales	Manual, service
India	3.426	3.092	10.310	3.793	1.033	1.475	0.994
China	1.170	0.697	3.140	0.764	1.205	0.633	0.515
Mexico	1.163	0.812	0.743	0.614	0.891	1.001	3.189
Philippines	0.931	0.513	0.895	3.321	0.380	0.983	1.479
South Korea	0.803	0.702	0.945	0.943	0.891	0.804	0.661
Vietnam	0.626	0.420	1.290	1.329	0.247	0.481	0.596
Canada	0.541	0.648	0.669	1.001	0.543	0.391	0.212
Taiwan	0.378	0.361	0.774	0.489	0.326	0.254	0.115
Pakistan	0.318	0.271	0.452	0.842	0.082	0.372	0.239
Colombia	0.308	0.272	0.247	0.366	0.301	0.296	0.466
United Kingdom	0.299	0.437	0.381	0.252	0.303	0.191	0.101
Germany	0.286	0.345	0.405	0.203	0.316	0.213	0.127

Notes: Each value represents the percentage share of a national-origin group in total hours worked by 25 to 54 year old males with at least four years of college in the 2011-2013 ACS sample.

The above occupation categories, which are based on Hanson and Liu (2017), are defined as follows:.

- Management, finance, and accounting: accountants, chief executives, financial managers, general managers, market surveyors, and economists.
- STEM: architects, computer programmers and software developers, engineers, life and medical scientists, mathematicians, and physical scientists.
- Health: dentists, pharmacists, physicians, registered nurses, therapists, and veterinarians.
- Education, law, social work, and the arts: instructors and teachers, lawyers, social and religious workers, writers, and artists.
- Technical, sales, and administrative support: administrative support staff, clerks and record keepers, sales representatives, sales supervisors, and technicians.
- Manual and service work: workers in agriculture, construction, hospitality, household service, and personal service ; machine operators and production workers; mechanics and repairers.

Table A.2: DOT and O*NET task intensities by occupation

	DOT Variables				ONET Variables			
Occupation	Abstract	Talk	Routine	Manual	Cognitive	Social	Routine	Manual
Engineers	0.90	0.54	0.73	0.51	0.92	0.39	0.60	0.27
Scientists, mathematicians	0.88	0.35	0.61	0.35	0.90	0.42	0.54	0.29
K-12 teachers	0.86	0.97	0.24	0.35	0.56	0.22	0.81	0.82
Post-secondary teachers	0.86	0.95	0.21	0.24	0.53	0.19	0.59	0.58
Accountants	0.86	0.66	0.54	0.12	0.93	0.95	0.56	0.20
Financial managers	0.85	0.76	0.28	0.14	0.94	0.79	0.85	0.47
Computer system analysts	0.85	0.78	0.33	0.34	0.63	0.70	0.53	0.32
Business analysts	0.85	0.68	0.33	0.31	0.61	0.38	0.66	0.32
Health professionals	0.84	0.83	0.64	0.60	0.63	0.41	0.81	0.93
Sales supervisors	0.82	0.73	0.27	0.42	0.81	0.42	0.85	0.72
Other managers	0.82	0.86	0.20	0.29	0.62	0.38	0.96	0.74
Admin. managers	0.82	0.74	0.30	0.42	0.63	0.48	0.88	0.57
Finance, insurance	0.79	0.76	0.38	0.17	0.85	0.76	0.74	0.60
Computer software developers	0.77	0.72	0.53	0.17	0.88	0.67	0.38	0.17
Chief executives	0.76	0.82	0.28	0.43	0.75	0.38	0.94	0.47
Health assistants	0.71	0.69	0.77	0.63	0.64	0.57	0.83	0.93
Lawyers	0.68	0.90	0.14	0.05	0.27	0.58	0.87	0.52
Technicians	0.63	0.57	0.75	0.59	0.68	0.64	0.45	0.53
Clergy, artists	0.63	0.75	0.35	0.37	0.28	0.30	0.74	0.55
Clerical, high skill	0.63	0.82	0.28	0.24	0.55	0.70	0.68	0.56
Construction	0.62	0.44	0.60	0.65	0.53	0.40	0.41	0.46
Clerical, medium skill	0.58	0.62	0.55	0.22	0.43	0.80	0.51	0.73
Mechanics, repairers	0.53	0.43	0.61	0.77	0.33	0.37	0.28	0.42
Services, medium skill	0.49	0.87	0.16	0.79	0.30	0.49	0.77	0.80
Salespersons	0.45	0.91	0.26	0.42	0.57	0.38	0.89	0.53
Clerical, low skill	0.38	0.58	0.45	0.35	0.48	0.73	0.41	0.61
Services, low skill	0.37	0.68	0.34	0.55	0.23	0.41	0.44	0.72
Machine operators	0.33	0.24	0.63	0.60	0.37	0.50	0.14	0.23
Transportation	0.25	0.38	0.31	0.83	0.27	0.37	0.25	0.46

Notes: DOT and O*NET variables are first transformed to percentile rankings for each detailed SOC occupation code. We then we compute the average percentile ranking for each aggregate occupation code, weighting by total labor hours worked by prime-age, college-educated males.

B Constructing the test statistics for conditions (A′)-(C′)

Test statistic for condition (C′): We apply the multivariate delta method to estimate the asymptotic normal distribution of $\beta_{cog}\gamma_{com} - \beta_{com}\gamma_{cog}$ and construct the test statistic. First, let $\Psi(\Lambda) = \beta_{cog}\gamma_{com} - \beta_{com}\gamma_{cog}$, which is a function of $R^4 \rightarrow R$ and $\Lambda = (\beta_{cog}, \beta_{com}, \gamma_{cog}, \gamma_{com})$. The delta method gives the following asymptotic normal distribution:

$$\sqrt{n}\big(\Psi(\widehat{\Lambda}) - \Psi(\Lambda)\big) \longrightarrow_d \mathcal{N}(0, G(\Lambda)\Sigma G(\Lambda)').$$

where $G(\Lambda) = \big(\frac{\partial\Psi(\Lambda)}{\partial\beta_{cog}}, \frac{\partial\Psi(\Lambda)}{\partial\beta_{com}}, \frac{\partial\Psi(\Lambda)}{\partial\gamma_{cog}}, \frac{\partial\Psi(\Lambda)}{\partial\gamma_{com}}\big) = \big(\gamma_{com}, -\gamma_{cog}, -\beta_{com}, \beta_{cog}\big)$. Σ is the covariance matrix of Λ,

$$\Sigma \equiv \begin{bmatrix} \text{Var}(\beta_{cog}) & \cdots & \cdots & \cdots \\ \text{Cov}(\beta_{cog}, \beta_{com}) & \text{Var}(\beta_{com}) & \cdots & \vdots \\ \text{Cov}(\beta_{cog}, \gamma_{cog}) & \text{Cov}(\beta_{com}, \gamma_{cog}) & \text{Var}(\gamma_{cog}) & \vdots \\ \text{Cov}(\beta_{cog}, \gamma_{com}) & \text{Cov}(\beta_{com}, \gamma_{com}) & \text{Cov}(\gamma_{cog}, \gamma_{com}) & \text{Var}(\gamma_{com}) \end{bmatrix}.$$

The estimate of $G(\Lambda)$ is, $(\widehat{\gamma}_{com}, -\widehat{\gamma}_{cog}, -\widehat{\beta}_{com}, \widehat{\beta}_{cog})$ which is available from Table 2. Σ is estimated from the same OLS regression. A consistent estimate of the asymptotic variance for $\beta_{cog}\gamma_{com} - \beta_{com}\gamma_{cog}$ is,

$$\widehat{\text{Avar}}\big(\beta_{cog}\gamma_{com} - \beta_{com}\gamma_{cog}\big) \equiv G(\widehat{\Lambda})\widehat{\Sigma}G(\widehat{\Lambda})'.$$

Note that the asymptotic variance for $\beta_{cog}\gamma_{com} - \beta_{com}\gamma_{cog}$ also depends on the off-diagonal element in Σ, which are the co-variance among OLS coefficients (which could make our variance large).

Under a null hypothesis $\beta_{cog}\gamma_{com} - \beta_{com}\gamma_{cog} = 0$, the test statistic is calculated as,

$$\frac{\widehat{\beta}_{cog}\widehat{\gamma}_{com} - \widehat{\beta}_{com}\widehat{\gamma}_{cog}}{\sqrt{\widehat{\text{Avar}}\big(\Psi(\widehat{\Lambda})\big)}}$$

Under the null, the test statistic asymptotically approximates a standard normal distribution.

Joint Test Statistic for Condition (A′), (B′), (C′): Combining the central limit theorem and the multivariate delta method, we have the following:

$$\sqrt{n}(\widehat{\Theta} - \Theta) \longrightarrow_d \mathcal{N}\Bigg(\mathbf{0}, S\Bigg)$$

where $\widehat{\Theta}$ is a 3-by-1 vector of estimates with each element being $\widehat{\beta}_{cog}$,$\widehat{\gamma}_{com}$, $\widehat{\beta}_{cog}\widehat{\gamma}_{com} - \widehat{\beta}_{com}\widehat{\gamma}_{cog}$. Θ is the vector of true parameter values. S is the 3-by-3 asymptotic covariance matrix.

We construct the χ^2-test statistic in two steps. First, we estimate S using a bootstrap procedure. We draw 200 bootstrap samples from the original sample, and obtain 200 estimates for $\widehat{\Theta}$. We compute $\widehat{S}$ based on the 200 estimates. Second, we calculate the test statistic as,

$$\widehat{\Theta}'\widehat{S}^{-1}\widehat{\Theta} \longrightarrow_d \chi^2(3).$$

C Additional Tables

Table C.1: Complete OLS regression results for Table 2

	(1) Average math	(2) 75ptl math	(3) 90ptl math
$Cog_s \times cog_o$	8.663	8.580	10.46
	(2.641)	(2.215)	(2.155)
$Cog_s \times rou_o$	2.748	1.030	2.910
	(2.458)	(2.238)	(2.330)
$Cog_s \times man_o$	-5.620	-5.716	-5.991
	(1.419)	(1.423)	(1.404)
$Cog_s \times com_o$	1.975	1.620	2.338
	(2.154)	(2.345)	(2.213)
$Ling_s \times cog_o$	0.752	0.738	0.745
	(0.379)	(0.427)	(0.402)
$Ling_s \times rou_o$	-0.821	-0.828	-0.776
	(0.548)	(0.539)	(0.529)
$Ling_s \times man_o$	0.0942	0.0650	0.0612
	(0.367)	(0.364)	(0.357)
$Ling_s \times com_o$	0.956	0.966	0.900
	(0.446)	(0.454)	(0.444)
$Dist_s \times cog_o$	0.719	0.442	0.519
	(0.336)	(0.322)	(0.318)
$Dist_s \times rou_o$	-0.0988	0.0495	-0.131
	(0.291)	(0.309)	(0.309)
$Dist_s \times man_o$	-0.541	-0.540	-0.485
	(0.216)	(0.230)	(0.216)
$Dist_s \times com_o$	-0.489	-0.383	-0.520
	(0.290)	(0.318)	(0.312)
Observations	1809	1597	1751
Adjusted R^2	0.311	0.318	0.313
Number of countries	69	61	67
Summary statistics for $\log \Pi^s_{d,o}$			
Mean	Standard deviation	25ptl	75ptl
-0.18	0.86	-0.69	0.35

Notes: Regression units are 29 occupations × 61-69 countries. The sample is foreign-born males, 25-54 years old, with a college degree, and with positive earnings. Occupation and country of origin fixed effects are included in all regressions. Standard errors (in parentheses) are clustered by origin country.

Table C.2: PPML results for $\log \Pi^{s}_{d,o}$ using DOT task intensities, all US immigrants

	(1) Average math	(2) 75ptl math	(3) 90ptl math
$Cog_s \times cog_o$	12.01	11.38	13.04
	(2.787)	(2.852)	(2.739)
$Cog_s \times rou_o$	1.389	-0.545	1.418
	(2.968)	(3.063)	(2.908)
$Cog_s \times man_o$	-5.869	-5.857	-6.536
	(2.588)	(2.658)	(2.546)
$Cog_s \times com_o$	-1.827	-1.638	-1.300
	(3.041)	(3.089)	(2.974)
$Ling_s \times cog_o$	0.487	0.566	0.461
	(0.830)	(0.828)	(0.827)
$Ling_s \times rou_o$	-0.778	-0.849	-0.768
	(0.884)	(0.894)	(0.887)
$Ling_s \times man_o$	-0.303	-0.320	-0.331
	(0.759)	(0.763)	(0.759)
$Ling_s \times com_o$	0.953	0.904	0.861
	(0.910)	(0.910)	(0.909)
$Dist_s \times cog_o$	0.544	0.414	0.411
	(0.423)	(0.447)	(0.438)
$Dist_s \times rou_o$	-0.266	-0.104	-0.343
	(0.445)	(0.476)	(0.460)
$Dist_s \times rou_o$	-0.406	-0.461	-0.285
	(0.389)	(0.415)	(0.403)
$Dist_s \times com_o$	-0.155	-0.136	-0.205
	(0.459)	(0.485)	(0.474)
Observations	2001	1769	1943
Pseudo R^2	0.187	0.191	0.188
Number of countries	69	61	67

Notes: Regression units are 29 occupations $\times$ 61-69 countries. The sample is foreign-born males, 25-54 years old, with a college degree, and with positive earnings. Occupation and country of origin fixed effects are included in all regressions. Standard errors (in parentheses) are clustered by origin country.

Table C.3: OLS results for $\log \Pi^s_{d,o}$ using O*NET task intensities, all US immigrants

	(1) Average math	(2) 75ptl math	(3) 90ptl math
$Cog_s \times cog_o$	4.565	4.237	5.955
	(2.316)	(1.991)	(2.004)
$Cog_s \times rou_o$	-0.207	-0.815	-0.673
	(1.547)	(1.613)	(1.545)
$Cog_s \times man_o$	-8.861	-9.795	-8.828
	(2.164)	(2.085)	(2.133)
$Cog_s \times com_o$	12.76	13.78	12.75
	(3.051)	(3.344)	(3.010)
$Ling_s \times cog_o$	-0.177	-0.179	-0.132
	(0.481)	(0.462)	(0.465)
$Ling_s \times rou_o$	-0.211	-0.178	-0.156
	(0.405)	(0.406)	(0.401)
$Ling_s \times man_o$	-0.262	-0.456	-0.321
	(0.543)	(0.553)	(0.568)
$Ling_s \times com_o$	0.874	1.197	0.972
	(0.803)	(0.837)	(0.837)
$Dist_s \times cog_o$	1.149	1.101	1.037
	(0.313)	(0.307)	(0.293)
$Dist_s \times rou_o$	-0.225	-0.158	-0.198
	(0.205)	(0.221)	(0.215)
$Dist_s \times man_o$	-0.404	-0.276	-0.329
	(0.224)	(0.229)	(0.227)
$Dist_s \times com_o$	0.202	0.0898	0.148
	(0.442)	(0.481)	(0.456)
Observations	1894	1673	1836
Adjusted R^2	0.393	0.411	0.395
Number of countries	69	61	67

Notes: Regression units are 29 occupations $\times$ 61-69 countries. The sample is foreign-born males, 25-54 years old, with a college degree, and with positive earnings. Occupation and country of origin fixed effects are included in all regressions. Standard errors (in parentheses) are clustered by origin country.

Table C.4: OLS regression results for $\log \Pi^s_{d,o}$ using DOT task intensities, Canadian immigrants

	(1) Average math	(2) 75ptl math	(3) 90ptl math
$Cog_s \times cog_o$	9.241	8.573	8.752
	(3.147)	(2.936)	(2.891)
$Cog_s \times rou_o$	-5.750	-5.681	-5.774
	(2.795)	(2.567)	(2.543)
$Cog_s \times man_o$	4.251	3.435	3.281
	(3.379)	(2.960)	(2.935)
$Cog_s \times com_o$	0.364	0.649	0.818
	(4.045)	(3.802)	(3.939)
$Ling_s \times cog_o$	3.205	3.192	3.154
	(0.660)	(0.651)	(0.648)
$Ling_s \times rou_o$	-2.082	-2.067	-2.046
	(0.495)	(0.498)	(0.498)
$Ling_s \times man_o$	0.696	0.703	0.695
	(0.702)	(0.707)	(0.713)
$Ling_s \times com_o$	-0.147	-0.161	-0.171
	(0.706)	(0.711)	(0.721)
$Dist_s \times cog_o$	0.743	0.707	0.688
	(0.324)	(0.319)	(0.311)
$Dist_s \times rou_o$	0.366	0.401	0.413
	(0.264)	(0.270)	(0.271)
$Dist_s \times man_o$	0.0439	0.0466	0.0479
	(0.196)	(0.194)	(0.195)
$Dist_s \times com_o$	0.178	0.165	0.157
	(0.276)	(0.278)	(0.281)
Observations	555	555	555
Adjusted R^2	0.717	0.718	0.719
Number of regions	22	22	22

Notes: Regression units are 27 occupations $\times$ 22 origin regions. The sample is restricted to foreign-born males, 25-54 years old, with a college education, and with positive earnings. Linguistic proximity and log distance are measured between Canada and the origin region. Occupation and region-specific fixed effects are included in all regressions. Standard errors (in parentheses) are clustered by region.

Table C.5: OLS results for $\log \Pi^s_{d,o}$ using DOT task intensities, by immigrant age of arrival

	(1) Average math	(2) 75ptl math	(3) 90ptl math
	Immigrants arriving before age 13		
$Cog_s \times cog_o$	0.566	-0.129	2.136
	(3.000)	(2.634)	(2.753)
$Cog_s \times com_o$	0.896	2.779	2.337
	(3.388)	(3.435)	(3.209)
$Ling_s \times cog_o$	-0.959	-1.103	-1.036
	(0.623)	(0.640)	(0.640)
$Ling_s \times com_o$	1.654	1.665	1.538
	(1.058)	(1.099)	(1.096)
Observations	1211	1071	1158
Adjusted R^2	0.285	0.292	0.297
Number of Countries	68	60	66

Notes: Regression units are 29 occupations $\times$ 61-69 countries. The sample is foreign-born males, 25-54 years old, with a college degree, and with positive earnings. Occupation and country of origin fixed effects are included in all regressions. Standard errors (in parentheses) are clustered by origin country.

Table C.6: OLS results for $\log \Pi^s_{d,o}$ using O*NET task intensities, by immigrant age of arrival and native-born ancestry

	(1) Average math	(2) 75ptl math	(3) 90ptl math
Panel I: Immigrants arriving age 18+			
$Cog_s \times cog_o$	5.250	5.238	7.329
	(2.725)	(2.213)	(2.160)
$Cog_s \times com_o$	9.948	9.708	9.934
	(2.860)	(3.085)	(2.906)
$Cog_s \times cog_o$	-0.508	-0.515	-0.507
	(0.533)	(0.498)	(0.508)
$Ling_s \times com_o$	1.820	1.956	1.882
	(0.771)	(0.803)	(0.790)
Observations	1691	1496	1633
Adjusted R^2	0.359	0.364	0.359
Number of Countries	69	61	67
Panel II: Immigrants arriving before age 18			
$Cog_s \times cog_o$	-0.107	0.522	1.376
	(2.187)	(1.863)	(1.874)
$Cog_s \times com_o$	3.960	5.283	5.208
	(2.680)	(2.390)	(2.367)
$Cog_s \times cog_o$	-0.411	-0.467	-0.386
	(0.500)	(0.490)	(0.506)
$Ling_s \times com_o$	1.350	1.585	1.415
	(0.700)	(0.698)	(0.698)
Observations	1350	1193	1294
Adjusted R^2	0.255	0.249	0.262
Number of Countries	68	60	66
Panel III: Native born by region of ancestry			
$Cog_s \times cog_o$	1.279	3.029	1.781
	(2.011)	(1.623)	(1.912)
$Cog_s \times com_o$	2.907	2.811	5.493
	(3.099)	(2.622)	(2.980)
$Cog_s \times cog_o$	-0.473	-0.422	-0.568
	(0.250)	(0.177)	(0.245)
$Ling_s \times com_o$	-1.016	-1.116	-1.003
	(0.732)	(0.640)	(0.682)
Observations	795	714	743
Adjusted R^2	0.239	0.273	0.238
Number of Countries	28	25	26

Notes: Standard errors (in parentheses) are clustered by origin country.

Table C.7: OLS results $\log \Pi^s_{d,o}$ using DOT task intensities, immigrants with 7+ years in US

	(1) Average math	(2) 75ptl math	(3) 90ptl math
$Cog_s \times cog_o$	8.617	9.804	10.65
	(3.094)	(2.532)	(2.404)
$Cog_s \times rou_o$	3.599	2.502	3.396
	(2.873)	(2.787)	(2.498)
$Cog_s \times man_o$	-3.656	-3.435	-4.112
	(2.761)	(2.909)	(2.909)
$Cog_s \times com_o$	3.582	2.463	3.513
	(2.475)	(2.450)	(2.597)
$Ling_s \times cog_o$	0.996	1.092	0.982
	(0.733)	(0.777)	(0.771)
$Ling_s \times rou_o$	-0.500	-0.587	-0.469
	(0.711)	(0.680)	(0.676)
$Ling_s \times man_o$	0.209	0.260	0.154
	(0.784)	(0.780)	(0.777)
$Ling_s \times com_o$	1.182	1.059	1.149
	(0.538)	(0.539)	(0.539)
$Dist_s \times cog_o$	0.815	0.579	0.568
	(0.437)	(0.439)	(0.427)
$Dist_s \times rou_o$	-0.278	-0.302	-0.342
	(0.320)	(0.354)	(0.341)
$Dist_s \times man_o$	-0.427	-0.455	-0.376
	(0.455)	(0.500)	(0.493)
$Dist_s \times com_o$	-0.972	-1.149	-1.017
	(0.432)	(0.486)	(0.460)
Observations	1607	1426	1549
Adjusted R^2	0.337	0.334	0.338
Number of countries	69	61	67

Notes: Regression units are 29 occupations × 61-69 countries. The sample is foreign-born males, 25-54 years old, with a college degree, with positive earnings, and with seven or more years of residence in the US. Occupation and country of origin fixed effects are included in all regressions. Standard errors (in parentheses) are clustered by origin country.

Table C.8: OLS results for $\log \Pi^s_{d,o}$ using DOT task intensities, immigrants arriving in US at age 18+ and with 7+ years of US residence

	(1) Average math	(2) 75ptl math	(3) 90ptl math
$Cog_s \times cog_o$	10.62	11.03	12.72
	(3.063)	(2.564)	(2.393)
$Cog_s \times rou_o$	2.654	1.127	2.907
	(2.670)	(2.562)	(2.412)
$Cog_s \times man_o$	-4.862	-3.843	-4.965
	(2.253)	(2.418)	(2.431)
$Cog_s \times com_o$	2.112	1.461	2.346
	(2.364)	(2.722)	(2.469)
$Ling_s \times cog_o$	0.481	0.530	0.470
	(0.510)	(0.566)	(0.544)
$Ling_s \times rou_o$	-0.869	-0.930	-0.858
	(0.606)	(0.572)	(0.571)
$Ling_s \times man_o$	-0.320	-0.313	-0.367
	(0.550)	(0.549)	(0.547)
$Ling_s \times com_o$	1.062	1.024	1.026
	(0.454)	(0.468)	(0.453)
$Dist_s \times cog_o$	0.511	0.193	0.251
	(0.390)	(0.380)	(0.366)
$Dist_s \times rou_o$	-0.338	-0.275	-0.408
	(0.293)	(0.326)	(0.310)
$Dist_s \times man_o$	-0.711	-0.793	-0.669
	(0.353)	(0.391)	(0.389)
$Dist_s \times com_o$	-0.989	-0.995	-1.022
	(0.410)	(0.470)	(0.442)
Observations	1691	1496	1633
Adjusted R^2	0.355	0.358	0.358
Number of countries	69	61	67

Notes: Regression units are 29 occupations $\times$ 61-69 countries. The sample is foreign-born males, 25-54 years old, with a college degree, with positive earnings, and with seven or more years of residence in the US. Occupation and country of origin fixed effects are included in all regressions. Standard errors (in parentheses) are clustered by origin country.

Table C.9: OLS results for $\log \Pi^s_{d,o} - \log \Pi^s_{d,o'}$ on occupational downgrading

Occupations	Comp-Engin	STEM	Non-STEM
Degree of Field	Comp-Engin	STEM	STEM
	Panel A: All Immigrants		
Cog_s	8.254	8.938	5.551
	(1.945)	(1.937)	(2.075)
$Ling_s$	1.102	0.860	0.316
	(0.612)	(0.550)	(0.583)
$Dist_s$	0.795	0.788	0.692
	(0.313)	(0.292)	(0.308)
Observations	63	65	61
	Panel B: Immigrants arriving age 18+		
Cog_s	7.641	8.187	4.978
	(2.422)	(2.367)	(2.480)
$Ling_s$	0.908	1.083	0.521
	(0.890)	(0.663)	(0.683)
$Dist_s$	0.758	0.955	0.721
	(0.432)	(0.351)	(0.364)
Observations	60	62	56
	Panel B: Immigrants arriving age 18+ with at least 7 years in the US		
Cog_s	7.514	8.038	4.873
	(2.527)	(2.366)	(2.378)
$Ling_s$	0.470	0.138	-0.0501
	(0.923)	(0.740)	(0.740)
$Dist_s$	0.755	0.660	0.414
	(0.444)	(0.377)	(0.377)
Observations	58	59	54

Notes: The dependent variable is $\log \Pi^s_{d,o} - \log \Pi^s_{d,o'}$, where o is the occupation listed in the first row, and o' is the set of low skill-occupations (construction, transportation, machine operation, mechanics and repairers, low-skill services, mid-skill services). The sample of immigrants is restricted to those with primary or secondary degree fields in computer-engineering subjects, in column (1), and STEM subjects, in columns (2) and (3). The number of observations equals the number of countries in the sample.

Table C.10: Fraction of students by proficiency level in the PISA math exam, 2015

	Level 0	**Level 1**	**Level 2**	**Level 3**	**Level 4**	**Level 5**	**Level 6**
Mean	0.153	0.181	0.221	0.209	0.147	0.067	0.021
Standard deviation	0.139	0.075	0.041	0.069	0.079	0.054	0.027

Notes: Summary statistics are for 58 countries in 2015 from the National Center for Education Statistics.

Table C.11: Changes in occupational wages and total efficiency units of labor (%), perfect substitution between foreign and native-born workers

	Wage units	Total efficiency units
Accountants	-0.038	-0.045
Admin. managers	0.016	-0.133
Business analysts	-0.023	-0.069
Clergy, artists	0.039	-0.168
Clerical, high skill	0.019	-0.136
Clerical, low skill	0.014	-0.128
Clerical, medium skill	0.011	-0.124
Computer software developers	-0.191	0.200
Computer system analysts	-0.037	-0.047
Construction	0.009	-0.121
Engineers	-0.085	0.030
Finance, insurance	-0.022	-0.071
Health assistants	0.013	-0.127
Health professionals	0.035	-0.161
Lawyers	-0.032	-0.055
Machine operators	0.005	-0.114
Chief executives	-0.020	-0.074
Financial managers	0.016	-0.132
Other managers	0.029	-0.152
Mechanics, repairers	0.011	-0.124
Sales supervisors	0.002	-0.110
Salespersons	-0.041	-0.041
Scientists, mathematicians	-0.336	0.433
Services, low skill	0.038	-0.167
Services, medium skill	0.026	-0.147
K-12 Teachers	0.084	-0.240
Post-secondary teachers	-0.142	0.121
Technicians	0.002	-0.110
Transportation	0.010	-0.122

Notes: All values are in percentage terms for changes in wages per efficiency unit of labor and changes total efficiency units of labor, each at the occupation level. The reported values are calculated as $(\widehat{w}_{d,o} - 1) \times 100$ and $(\widehat{L}_{d,o} - 1) \times 100$.

Table C.12: Changes in occupational wages and total efficiency units of labor (%), imperfect substitution between foreign nad native-born workers

	Wage Units		**Total Efficiency Units**	
	Native	Immigrants	Native	Immigrants
Accountants	-0.023	0.123	0.001	-0.668
Admin. manager	-0.024	0.199	-0.004	-1.023
Business analysts	-0.021	0.041	0.008	-0.276
Clergy, artists	-0.029	0.434	-0.018	-2.121
Clerical, high skill	-0.025	0.359	-0.006	-1.756
Clerical, low skill	-0.023	0.134	0.001	-0.715
Clerical, medium skill	-0.023	0.213	-0.001	-1.079
Computer software developers	0.007	-0.383	0.090	1.902
Computer system analysts	-0.015	-0.121	0.025	0.516
Construction	-0.022	0.073	0.004	-0.432
Engineers	-0.017	-0.051	0.017	0.170
Finance, insurance	-0.023	0.161	0.002	-0.841
Health assistants	-0.023	0.156	-0.001	-0.821
Health professionals	-0.027	0.180	-0.011	-0.957
Lawyers	-0.021	0.192	0.006	-0.971
Machine operators	-0.021	0.071	0.005	-0.417
Chief executives	-0.023	0.141	0.001	-0.750
Financial managers	-0.027	0.292	-0.011	-1.464
Other managers	-0.027	0.423	-0.012	-2.057
Mechanics, repairers	-0.022	0.090	0.003	-0.511
Sales supervisors	-0.022	0.147	0.002	-0.772
Salespersons	-0.019	0.011	0.012	-0.128
Scientists, mathematicians	0.003	-0.430	0.077	2.095
Services, low skill	-0.025	0.124	-0.006	-0.687
Services, medium skill	-0.025	0.391	-0.005	-1.897
K-12 teachers	-0.032	0.974	-0.027	-4.529
Post-secondary teachers	-0.018	-0.046	0.017	0.151
Technicians	-0.022	0.156	0.001	-0.818
Transportation	-0.022	0.076	0.003	-0.443

Notes: All values are in percentage terms for changes in wages per efficiency unit of labor and changes total efficiency units of labor, each at the occupation level. The reported values are calculated as $(\widehat{w}_{d,o} - 1) \times 100$ and $(\widehat{L}_{d,o} - 1) \times 100$.

D Appendix - Counterfactual Exercise

D.1 Synthetic Control for India

To construct the synthetic control for India, we target 58 moments of the data that consist of India's occupation employment shares and average wages across the 29 occupations in US data. We choose the weights W that minimize the Euclidean difference,

$$||X_o^{India} - X_o W|| \tag{17}$$

where X_o^{India} is a 58×1 vector with the first 29 elements being the occupation employment shares of college-educated immigrants from India in the US, and the next 29 elements are the average occupational wages of college-educated immigrants from India in the US. X_o is a $58 \times n$ matrix, where n is the number of immigrant countries of origin used to generate the synthetic control. The first 29 rows of X_o consist of on occupation employment shares; the next 29 rows consist of average occupational wages. W has dimension n.

In order to find precise estimates of the weights, the number of targeted moments has to be much greater than n. We perform the synthetic control method using the 15 countries that have the highest PISA scores.[48] Among the 15 countries, six have positive weights and nine have zero weights.[49] The imputed PISA score density for India yields shares across the seven proficiency levels, from 0 to 6 respectively, of 4.1%, 10.8%, 20.2%, 26.0%, 22.4%, 12.0%, and 4.6%. The imputed share of scores in the top three categories of 39.0% for India is comparable to that of Belgium, Canada and the Netherlands.

D.2 Imperfect Immigrant-Native Substitution at Occupational Levels

We extend the model to allow imperfect immigrant-native substitution at occupational level. The aggregate efficiency units of labor in occupation o is

$$L_{d,o} = \left[\alpha_{d,o,n} N_{d,o}^{\frac{\lambda-1}{\lambda}} + \alpha_{d,o,f} F_{d,o}^{\frac{\lambda-1}{\lambda}}\right]^{\frac{\lambda}{\lambda-1}}.$$

$N_{d,o}$ and $F_{d,o}$ are the aggregate efficiency units in occupation o for native and immigrant workers, respectively. λ denotes the elasticity of substitution between immigrants and natives within occupation, which we set to be 4.6 following Burstein et al. (2020). We denote the wage efficiency unit per labor as $w_{d,o,n}$ for natives and $w_{d,o,f}$ for immigrants. The efficiency units of native and immigrant labor demand at each market are represented by the following:

$$N_{d,o}^{\text{demand}} = \frac{1}{w_{d,o,n}^{\lambda}} \frac{1}{p_{d,o}^{\rho}} Y_d A_{d,o}^{\rho} \alpha_{d,o,n}^{\lambda}$$

and

$$F_{d,o}^{\text{demand}} = \frac{1}{w_{d,o,f}^{\lambda}} \frac{1}{p_{d,o}^{\rho}} Y_d A_{d,o}^{\rho} \alpha_{d,o,f}^{\lambda},$$

where the price per unit of occupational output is

$$p_{d,o} = \left[\alpha_{d,o,n}^{\frac{1}{\lambda}} w_{d,o,n}^{1-\lambda} + \alpha_{d,o,f}^{\frac{1}{\lambda}} w_{d,o,f}^{1-\lambda}\right]^{\frac{1}{1-\lambda}}$$

In equilibrium, $w_{d,o,n}$ clears the market such that $N_{d,o}^{\text{demand}} = N_{d,o}^{\text{supply}}$, and $w_{d,o,f}$ clears the market such that $F_{d,o}^{\text{demand}} = F_{d,o}^{\text{supply}}$.

[48] These countries are Singapore, Taiwan, Hong Kong, China, Korea, Japan, Switzerland, Belgium, Netherlands, Canada, Estonia, Germany, Austria, Poland, and Finland.

[49] The weights are 0.11 for Singapore, 0.04 for Korea, 0.02 for Switzerland, 0.39 for Canada, 0.22 for Germany, and 0.21 for Poland.