Journal of Population Economics (2022) 35:89–133
https://doi.org/10.1007/s00148-020-00811-4

ORIGINAL PAPER

Wage distributions in origin societies and occupational choices of immigrant generations in the USA

Crystal Zhan[1]

Received: 8 April 2019 / Accepted: 18 November 2020 / Published online: 6 January 2021

Abstract

This paper studies occupational selection among generations of immigrants in the USA and links their choices to the occupational wage distribution in their country of origin. The empirical results suggest that individuals are more likely to take up an occupation in the USA that was more lucrative in the origin country, conditional on individual demographic characteristics, parental human capital, and ethnic networks. However, the importance of the origin wage declines with the length of time that immigrants spend in the USA and over generations. Information frictions may be an explanation.

Keywords Immigrants · Occupational choices · Origin wage · Intergenerational transmission · Assimilation · Human capital · Information

JEL Classification J15 · J24 · J31

1 Introduction

Immigrants to the USA are a growing and increasingly diverse population. The number of immigrants in the USA quadrupled from 1970 to 2013, and the origin of immigrant inflows has shifted away from Europe to Latin America and Asia. As the foreign-born population has proliferated in the USA, so has that of second-generation immigrants (Blau et al. 2013).[1]

[1] Second generation refers to individuals who were born in the USA but have at least one foreign-born parent.

Responsible editor: Klaus F. Zimmermann

✉ Crystal Zhan
Crystal.Zhan@moore.sc.edu

[1] Department of Economics, Darla Moore School of Business, University of South Carolina, 1014 Greene Street, Columbia, SC 29208, USA

Changes in the scope and composition of immigration have brought fundamental changes to the US labor market. Evidence suggests that the economic outcomes of immigrants and their descendants in a host society vary by country of origin and are related to the characteristics of those countries (Blau et al. 2011, 2013; Abramitzky and Boustan 2017; Hanson and Liu 2017). Influences from source countries are then transmitted across generations and last for decades.

In this paper, I investigate whether occupational wage distributions in origin countries influence the job selection of generations of immigrants in the USA. The choice of occupation is one of the most important economic decisions an individual makes. Occupations determine earnings and socioeconomic standing (Constant and Zimmermann 2003). For immigrants, their assimilation progress in the host society hinges on their labor market performance.

The occupational wage structure is formed based on market demand and supply of labor, both of which are linked to many factors. Consequently, origin occupational wages can affect immigrants and their offspring in the USA through multiple channels. First, the origin wage distribution may capture discrepancies in unobserved human capital across national origins. Such discrepancies might result from education systems (Wobmann 2003), job specialization, and selective migration (Brjas 1987, 1995b). Due to intergenerational transmission of human capital and ethnic networks (Edin et al. 2003; Lafortune and Tessada 2012), such discrepancies can persist across generations. Second, some immigrants may view their origin country as a potential destination labor market and take the wages there into consideration when choosing jobs in the USA. Third, cultural norms may shape individuals' preferences toward specific occupations or occupational attributes (Bisin and Verdier 2000; Zhan 2015). These preferences may persist post-migration. Lastly, immigrants may have limited information on local labor markets in the host country and use their pre-migration experience to inform their US job selection.

While the reliance on the origin labor market may impede immigrants' assimilation to the US labor market, the four channels have different implications. The first two channels suggest that immigrants are sorting into the "right" occupations given unobserved factors that are correlated with origin occupational wages. The third channel also implies the selection of the "right" occupations according to endogenous preferences, which are formed on the basis of country-of-origin wages. The last channel, however, indicates that immigrants choose the "wrong" occupations by considering the wage in the origin country, an irrelevant attribute. If they were to ignore origin wages, immigrants could potentially get higher-paid jobs. Hence, the misinformation channel suggests a welfare loss, whereas the others do not.

In this paper, I assess three generations of immigrants surveyed in the 2005–2018 American Community Surveys: the first generation that migrated to the USA after the age of five, the 1.5 generation that was foreign-born but came to the USA at the age of five or less, and the second and higher generations that were native-born.[2] The information about occupational wages in source countries is gathered from the

[2]According to statistics from the Current Population Survey, a nontrivial fraction of the native-born Americans in the sample may be second-generation immigrants.

Occupational Wages around the World Database by Richard B. Freeman and Remco H. Oostendorp. I restrict the sample to prime-aged males with a bachelor's degree or higher in the main analysis. Using a highly educated sample reduces potential problems associated with self-selection into the labor force and unobserved ability differences (Dolton et al. 1989; Arcidiacono 2004). I also use general occupational categories so that the access to occupations is less constrained by a specialized skill set or location. As a result, the individuals in the sample are more likely to face a comparable set of occupation choices and have the discretion to select the most favorable one. While highly skilled workers comprise an important component of current immigrant inflows (Peri and Sparber 2011), limiting the sample to college graduates may pose issues to the internal and external validity of the study. In the extended analysis, therefore, my sample includes high school graduates and individuals with some college additionally.

I model occupational selection as a conditional logit discrete choice problem where an individual is assumed to choose the occupation that yields the highest utility. The hedonic utility is a function of the attributes of an occupation and its wage income in one's origin country. The regressions also control for factors related to opportunity or advantage, such as individual demographic characteristics, family background, and origin country characteristics that may affect selective migration. Some specifications include origin fixed effects to further eliminate the possibility that unobserved individual characteristics drive the occupational choices.

The empirical results suggest that all three generations place a positive weight on the country-of-origin occupational wage when choosing jobs in the US, conditional on the expected earnings in the US market. The positive effect of the origin wage persists across generations, though its scope shrinks. These results are robust to alternative specifications that address potential problems related to unobserved heterogeneity in human capital, the violation of the assumption of the independence of irrelevant alternatives, measurement errors in the US occupational wages, and the period to which the origin country's wage refers. Focusing on the first-generation immigrants, I find that the length of time in the USA mitigates the effect of the origin wage. The temporal change is small, however, which may suggest assimilation takes place at a slow pace.

Subsequently, I explore the potential mechanisms underlying these empirical patterns. I find that a higher origin wage for a person's chosen US occupation does not predict higher actual US earnings, controlling for the average wage of that occupation in the USA. Indeed, among the native-born generation, individuals' realized income is significantly and negatively related to the origin wage of their job. This finding may rule out the possibility that the origin occupational wage captures the unobserved abilities, motivation, or ethnic networks that lead to higher productivity.

I next use an alternative dataset, the Knowledge of the World of Work survey, to assess the information acquired by individuals of different ancestries. This survey asked young adults for their opinions on earnings of various occupations. These young adults might have little market experience at the time of the survey, but their ranking of occupational earnings conforms to the wage distribution in their origin countries. This pattern may confirm that information frictions are a plausible

explanation for the positive weight assigned to origin occupational wages when immigrants choose jobs in the USA.

Finally, I expand the sample to include individuals with a high school degree or some college. Because qualifications and ability may restrict the set of occupation choices available to less-educated workers, I employ a nested logit model and categorize the occupations into two nests: professional and non-professional. I assume that years of education and other individual characteristics capturing advantage and opportunity determine the choice of the nest, and the origin occupational wage affects the selection within a nest. The results verify that better-educated people are more likely to work in professional jobs. Conditional on the choice of the nest, individuals exhibit a high propensity to take up an occupation that was more profitable in the origin. As in the sample of college graduates, the weight placed on the origin wage dissipates over generations.

The findings in this paper underscore the importance of ethnic backgrounds in immigrants and their descendants' occupational choices in the USA. These findings may help policymakers formulate optimal immigration and social policies. For instance, if ethnic backgrounds guide immigrants to sort into specific occupations, different national-origin groups will contribute to the host economy differently. Such a trend may have implications for the type of immigrants that a host country chooses to target and how to create appropriate incentives to attract them (Chiswick and Taengnoi 2007). The results may also shed light on the potential impact of immigrants on native and previous immigrant workers. Policies to restrict entry to certain occupations or industries for new immigrants could prove helpful to protect the well-being of the other two groups (Lafortune and Tessada 2012).

2 Literature review and contribution

This paper belongs to a broad literature on economic outcomes and assimilation of immigrants and their offspring in receiving countries. The existing literature has established a set of patterns, estimating both the economic and cultural convergence between immigrants and natives (see, for example, Chiswick (1978), Borjas (1995a), Abramitzky and Boustan (2017)). The paper contributes to this literature by identifying an important social and economic factor—pre-migration experience, as captured by the country-of-origin wage distribution—to explain immigrants' labor market performance and the assimilation process in the USA. Underscoring the importance of information in one's host-country job selection, the paper adds to a series of studies that investigate the occupational choices of immigrants, across national origins and relative to natives (Chiswick and Taengnoi 2007; Peri and Sparber 2011; Hanson and Liu 2017). This paper verifies a slow pace of assimilation of immigrants and varying speeds of convergence across sending countries, both of which have been established by several previous papers (Chiswick 1983; Abramitzky et al. 2014). The results in this paper may imply that a more substantial disparity in the wage structure between the USA and the country of origin can result in a more severe skill-job mismatch and, as a result, a penalty on income. This paper, then, echoes multiple studies that inspect

the impact of ethnic networks, or ethnic capital, on immigrants' labor market outcomes (Borjas 1992; Munshi 2003; Lafortune and Tessada 2012). Ethnic capital may depend on the wage structure of an origin country, and social networks are pivotal to the transmission of information in the host country. Lastly, the paper is related to a set of papers on selective migration (Borjas 1987; Abramitzky and Boustan 2017; Spitzer and Zimran 2018). By linking immigrants' US earnings to country-of-origin occupational wages, the paper provides evidence of a positive selection of migrants from lower-paid jobs in the home country but a negative selection from higher-paid jobs, a pattern consistent with that proposed by Borjas (1987) .

Furthermore, the paper contributes to a second strand of literature on individual choices of college majors and occupations (see, for example, Boskin (1974), Arcidiacono (2004), Ryoo and Rosen (2004), Propper and Reenen (2010)). The majority of this literature focuses on expected lifetime earnings and job tasks and demonstrates that people respond to variation in salaries when making career choices. This paper adds to the literature by addressing the factors of ethnic backgrounds, past labor market experience, and information flows in job selection. These factors may impact not only international migrants but also people who move across local markets or encounter policy shocks.

3 Conceptual framework

This section discusses the mechanisms underlying the connection between the country-of-origin wage structure and immigrants' occupational choice in the USA. There are four possible channels: unobserved human capital, mobility, cultural preferences, and information.

1. The *unobserved discrepancies in human capital* across national origins may result in differential expected earnings of each occupation for individuals with different origins. Such discrepancies may be correlated with the wage distribution in one's country of origin for the following reasons.

    - *Ability differences*: The educational infrastructure and systems may result in more accumulation of certain skills in a country than in other countries (Wobmann 2003). Hence, the labor supply to the occupations where such skills can apply would be larger, and the wage (standing) of these occupations would be lower, conditional on the demand. Accordingly, the wage rate of an occupation in the origin country may negatively predict the probability of immigrants taking up that occupation in the host country.

      Conversely, a higher occupational wage in a country may imply more specialization in that occupation. Thus, the migrants from that country may be more productive to work in that occupation and receive higher US earnings than people from elsewhere. A positive connection between immigrants' US occupation and their origin-country wage is expected. Notably, the specialization premise may apply to individuals who have accumulated occupation-specific human capital before migration but not to those

who migrate at a young age or to the native-born children of immigrants, conditional on their parents' educational and occupational backgrounds.

- *Selective migration*: The wage distributions of sending and receiving countries may impact the selection of migrants, which takes place in two possible ways. First, individuals will move to the USA if they have a comparative advantage (relative to the natives). Second, they will migrate if the returns on their skills are higher in the USA than in their home country (Borjas 1987). In both scenarios, a lower (relative to the USA) wage standing in a country predicts more emigrants with the skills required by that occupation. If immigrants engage in the same or similar jobs post-migration, they are more likely to work in occupations associated with a lower country-of-origin wage. On the other hand, if migration is more likely among those who can afford the trip and have better English skills, it might be those who were paid well in their home countries that come to the USA. In this case, origin wages may positively predict one's job choice in the USA.
- *Intergenerational transmission of human capital*: Children of immigrants may engage in similar occupations as their parent generation, as they inherit skills and ability from their parents and ethnic communities (Borjas 1995b). This dynamic will lead to serial correlation in occupational choice across generations. So first-generation investments related to source-country earnings will persist through multiple generations.
- *Ethnic networks*: Ethnic networks may facilitate one's job search (Munshi 2003; Edin et al. 2003; Lafortune and Tessada 2012). If certain occupations in the home country are common choices for the previous generation, the descendants may have access to networks in those jobs that increase their expected earnings in the USA relative to individuals from other ethnic backgrounds. It is also possible that ethnic enclaves slow down the post-migration accumulation of USA-specific human capital and affect immigrants' earnings negatively (Warman 2007; Eriksson 2019). Both scenarios may generate positive occupational wage correlation across the two countries.

2. *Return migration* is possible for some immigrants. If individuals consider their country of origin as a future destination labor market, expected (lifetime) earnings of each occupation will depend on both USA and country-of-origin wages. This is especially true for first-generation immigrants, many of whom moved to the USA after making education and initial occupation choices in their country of origin. One-point-five-generation workers, who are likely fluent in their parent's native language as well as in English, may consider a return to their parents' source country as a viable possibility. They may also, therefore, place some weight on the country-of-origin wages when considering expected earnings. The implications for the native-born are less clear, although in the current data, a non-trivial fraction of them may have immigrant parents and, thus, form expectations similar to those of 1.5 generation.
3. *Social and cultural norms* in a country shape the preferences of its people. Immigrants may carry with them such preferences and ideas about occupations based on personal experiences and common social representations in source countries.

They may later socialize their US-born children with the same set of values (Bisin and Verdier 2000; Giuliano 2007). These norms and attitudes may predict the link between origin occupational wage and US job selection of immigrants in both directions. On the one hand, people often attach social prestige to high-income jobs. Hence, a lucrative job in the country of origin may also be highly sought-after in the USA regardless of the actual wage offered. On the other hand, when a society values an occupation culturally, a surplus labor supply reduces the compensating wage for that occupation. If migrants continue to favor that occupation in the USA, a negative link may be observed.

4. *Information frictions* can cause mistaken beliefs about which occupations are higher earning. The language barrier and lack of social networks can prevent new immigrants from getting the correct information about occupational earnings in local labor markets. These job-seekers may, thereby, view an occupation according to their pre-migration experience. Similarly, because of misinformation, immigrants' choices of human capital investment for their children may depend on judgments developed in their home countries. A positive connection between occupational wage in the country of origin and the occupational choices of immigrants to the USA can be expected.

The first two channels may imply that immigrants and their descendants make occupational selections in the USA based on correct but unobservable (to the econometrician) information. These channels are consistent with findings in the existing literature that occupational choice responds to expected earnings (e.g., Boskin 1974 and Arcidiacono 2004). Nevertheless, individuals who migrate at a young age or who are native-born to immigrant parents may have no direct exposure to the country-of-origin labor market. Disentangling unobserved human capital from other channels may be possible for these individuals by controlling for the education and occupations of their parents and the broader ethnic environment, including ethnic networks. The propensity of return migration is also lower for later generations. The third channel addresses the influence of a non-economic factor: cultural preference. Individuals may have correct information about US occupational wages, but their optimal choices do not necessarily yield the highest expected earnings because of unobserved preferences. The last channel, in contrast, indicates that immigrants use incorrect information to guide their occupational choices and, as a result, may suffer an income penalty. Immigrants who encounter higher obstacles to accessing information, such as those without competent English skills or those from countries whose economic structure differs dramatically from that of the USA, may be more adversely impacted. Indeed, these groups are precisely those who have more difficulties in assimilating to the host society. Individuals in this situation could potentially improve their labor market performance if policy intervention provides them with correct information.

4 Empirical model

I study the occupational choices of immigrants and their descendants in the following manner. Suppose individuals select their occupation from W mutually exclusive

options to maximize their utility. The hedonic utility for individual i of birth cohort t from origin j to work in occupation w, where $w = 1, 2, ..., W$, is:

$$U_{ijwt} = V\left(Y_{wi}, X_i, Z_{jwt}\right) + \xi_{ijwt}.$$

Y_{wi} measures the attributes of occupation w at the time when individual i chooses the current job that may determine expected lifetime earnings. These attributes may include wage income, training costs, working hours, and the likelihood of unemployment. X_i is a vector of observed individual characteristics that may be related to opportunity and advantage and affect job selection, such as age, working disability, and English proficiency. For first-generation immigrants, I also include the number of years an individual has lived in the USA, as immigrants may accumulate USA-specific human capital and information about labor markets over time. Their immigration status may change with the length of the migration spell as well. Additionally, I introduce year-of-survey fixed effects to X_i to account for market fluctuations across years. Z_{jwt} represents the wage of occupation w in country of origin j that may influence the job decisions of cohort t. Since the wage distribution in a market evolves, one may argue that origin wages before first-generation immigrants migrated to the USA would best reflect the origin-specific human capital, cultural norms, or information they carried with them and then transmitted to their descendants. However, one may also argue that the wage distributions faced by counterparts of the US population, or their "cousins" in origin countries, may better capture exposure and evolution of the origin traits that have influenced the immigrants (Fernandez and Fogli 2009; Zhan 2015). Data limitations discussed in the following paragraphs only allow me to match generations in the USA to their country-of-origin occupational wages based on the second argument. Besides, if mobility is the reason why individuals in the USA value the wages in their origin countries, it should be the current wage rates that matter.[3] $V\left(Y_{wi}, X_i, Z_{jwt}\right)$ can thereby be interpreted as individual i's evaluation of occupation w. The error term ξ_{ijw} represents unobservable preferences.

The probability that individual i selects occupation w, p_{ijwt}, is:

$$p_{ijwt} = \Pr\left(U_{ijwt} \geqslant U_{ijkt}\right), \forall k = 1, 2, ..., W. \tag{1}$$

The error term ξ_{ijw} is assumed to be independent across occupations for each individual. So the likelihood that a particular occupation is chosen over another one is independent of other alternatives. If the error terms follow a standard type I extreme value distribution, p_{ijwt} can be derived to be:

$$p_{ijwt} = \frac{\exp\left(V_{ijwt}\right)}{\sum_{k=1}^{W} \exp\left(V_{ijkt}\right)}, \tag{2}$$

where $V_{ijwt} = V\left(Y_{wi}, X_i, Z_{jwt}\right)$.

In a conditional logit model, the coefficients on variables that are not alternative-specific, such as individual characteristics, cannot be directly estimated. However,

[3] I conduct several sensitivity tests regarding the period to which the origin wage refers and provide the details of these tests in online appendix (https://drive.google.com/file/d/1gR-rPtQhcIdxpgRuHmAqu6obsXiAbnwz/view?usp=sharing).

the effects of such variables can be identified by interacting them with the attributes of the options.[4] That is, these factors are assumed to affect job selection through their influence on how individuals evaluate various occupational attributes. Therefore, a linear formulation of V_{ijwt} is:

$$V_{ijwt} = \eta_1' Y_{wi} + \eta_2' Y_{wi} \times X_i + \phi_1' Z_{jwt} + \phi_2' Z_{jwt} \times X_i + e_{ijwt}. \quad (3)$$

e_{ijwt} includes factors that are related to occupational choices but unobserved by the econometrician. ϕ_1 and ϕ_2 capture the effects of occupational wage distribution of origin countries on the occupational choices of individuals in the US market. If X_i is standardized to have a mean zero and standard deviation of one across the sample, a positive ϕ_1 suggests that an individual with mean characteristics in the sample would favor a job that offers high monetary reward in the country of origin, and vice versa.

To disentangle the intergenerational transmission of human capital and rule out the probability that occupational wages in origin countries pick up other systematic variation across origins, such as selective migration, I further include in X_i parental income, parental education, the number of siblings,[5] and several economic and political factors of the source country that may impact the emigration decisions of its people (i.e., the GDP per capita, Gini coefficient, average educational attainment, war, democracy, share of the labor force in agriculture, distance to the USA, and whether English is an official language).[6] Because a country's wage structure evolves, I also control for country fixed effects in some specifications by interacting country-of-origin dummies with occupational attributes, including the the origin wage Z_{jwt}. This way, the unobserved time-invariant discrepancies across countries can be effectively removed, and the identifying variation comes solely from temporal changes in the origin wages. Finally, because ethnic networks do not only facilitate individuals in finding employment but also captures the set of specialized human capital of an ethnic group, I include the proportion of the parental generation of individual i working in occupation w, l_{ijw}, as a measure of ethnic networks.

Hence, Eq. 3 is modified to:

$$V_{ijwt} = \eta_1' Y_{wi} + \eta_2' Y_{wi} \times X_i + \phi_1' Z_{jwt} + \phi_2' Z_{jwt} \times X_i + \lambda' l_{ijw} + \varepsilon_{ijwt}. \quad (4)$$

With the additional controls, ε_{ijwt} is less likely to correlate with origin occupational wage Z_{jwt}.

[4] Variables that do not vary across occupations may also enter the regression by interacting them with a set of occupation dummies. I use the interaction of person-specific characteristics and occupational attributes for computational advantage.

[5] Because the American Community Surveys (ACS) does not document respondents' parental characteristics, I estimate these characteristics using a group-mean method (Card et al. 2000; Blau et al. 2013). The next section reveals more details about this method.

[6] The last two time-invariant traits are dropped from the regression when country fixed effects are introduced.

Nevertheless, the variation in parental socioeconomic backgrounds and their occupational patterns may be partly attributed to the origin wage structure. Controlling for these variables may, therefore, result in underestimation of the impact of origin occupational wages.

5 Data description

5.1 Sample selection

The main dataset employed is the Integrated Public Use Microsample Series version of the 2005–2018 American Community Surveys (ACS). I focus on males aged 35–54 and exclude those living in group quarters. Individuals in this age range have usually completed their education, and differential mortality is unlikely to be a problem. Compared to males, female samples are more affected by selection bias due to labor force participation decisions. Female immigrants are also more often tied movers, whose labor market performance may result from compromised choices (Keith and McWilliams 1999). In the primary analysis, I restrict the sample to individuals with bachelor's degrees or higher to reduce problems associated with unobserved ability. The occupational pattern of low-skilled workers may be largely driven by their qualification for different jobs rather than their preferences. In comparison, the highly educated have a better chance to select occupations that maximize their utility. Additional years of postgraduate studies may be viewed as part of the career path. Nonetheless, lower-educated individuals make up an essential part of the US labor force, and so I include high school graduates and individuals with some college in the sample additionally in extended analysis.

I assign ethnic origins of individuals by their first-reported ancestry. For a foreign-born individual who does not report an ancestry, I define birthplace as the ethnic origin. To investigate the assimilation process, I divide the sample into three groups: the first generation, the 1.5 generation,[7] and native-born Americans who belong to second or higher generations. In the ACS, about 60% of native-born males reported an identifiable non-US ancestry. These people might retain a stronger ethnic identity or a closer link to their ethnic community than those who did not report and thus may not well represent the whole US-born population.[8] Moreover, I drop origin countries displaying less than 10 observations in each generation. Due to computational constraints of discrete choice models, for the origins with the number of observations

[7]Second-generation immigrants are not identifiable in the ACS data. Instead, I examine the 1.5-generation immigrants who migrated to the USA prior to the age of five. These people may obtain all their education in the USA but may be tightly connected to the origin society. Therefore, the 1.5 generation should have a close resemblance to the second generation.

[8]Because both the ACS and the Current Population Survey (CPS) survey nationally representative samples, and because second-generation immigrants can be distinguished in the CPS data, I compare the statistics from the ACS and the CPS of the same years. Were all second-generation individuals to identify their ancestries, statistics suggest that approximately 20% of the native-born American sample are second-generation.

that well exceeds 1500 in the first and last group, I randomly select around 1500 individuals from each ethnic origin.[9]

In addition to ancestry and nativity, the ACS data document individuals' educational attainment, employment, occupation, and other basic demographic characteristics such as age, gender, and working disability. Unfortunately, the ACS does not collect information on respondents' parents. Because intergenerational human capital transmission plays a critical part in one's occupational choice, I follow Card et al. (2000) and Blau et al. (2013) and use a group-mean method to estimate the parental characteristics. I identify the groups of "potential parents" of the 1.5 generation and higher generation from the earlier censuses[10] and derive the means of years of education, family income, and the number of children to proxy for the actual parental characteristics.[11] I also calculate the proportion of the matched parental generation working in each occupation to measure the ethnic network in the labor market. These proportions may serve as proxies for the occupations of one's own parents as well.

Table 1 reports the summary statistics by generations. Among the three groups, first-generation immigrants are the youngest and least likely to be disabled; they have the most postgraduate education but the lowest wage income. The 1.5 generation resembles the native-born in age and years of schooling. Yet the 1.5 generation has a higher share of disabled individuals, a slightly larger share of people who do not speak English, and a lower income than the later generation. The individuals in the sample were born between 1956 and 1985. For analytical purposes, I define cohorts by every five birth years.

5.2 Occupations

The occupation reported in the ACS is the job from which respondents earn the most money or the job at which they spend the most time. Unemployed persons or those out of the labor force for less than five years were to give their most recent jobs.[12] There are 383 occupations in total in the 1990 Census Bureau occupational classification, excluding the armed services. Highly educated workers engage in all of the 383 occupations.

I aggregate the three-digit-coded jobs into 23 categories based on the Census/ACS occupation grouping and the required training.[13] The aggregation reduces the

[9] I do not trim the sample of 1.5 generation as this group is much smaller. Appendix Table 11 reports the list of origin countries in the sample and the number of observations in each generation by origin (pre- and post-trimming). Alternatively, I form a 25% random sample of the first generation and a 5% random sample of the native-born Americans, regardless of countries of origin. These two samples have a comparable size to the samples analyzed in the paper. Repeating the regressions in Tables 5 and 6 on the random samples yields very similar results. The results are available from the author upon request.

[10] The parents of the 1.5 and higher generations are more likely to be present in the USA and be observed in the earlier censuses. It is not clear whether the parents of the first-generation immigrants migrated to the USA or not. Therefore, I do not estimate parental education, income, or the number of siblings for the first generation. Nevertheless, the parental ethnic networks estimated for the 1.5 generation may also reflect the ethnic networks of the first generation in the USA.

[11] Appendix Section A.1 offers more information about the group-mean method.

[12] In the ACS sample, about 4% are out of the labor force, and about 3% are unemployed.

[13] Occupation categorization is described in more detail in the Appendix Section A.2.

Table 1 Summary statistics on different generations

	First Gen.		1.5 Gen.		Native-born	
Variable	Mean	SD	Mean	SD	Mean	SD
lIndividual characteristics						
Age	43.4	5.22	43.2	5.27	43.3	5.33
Disability (= 1)	.022	.148	.031	.172	.028	.165
Speak English (= 1)	.925	.263	.991	.092	.996	.064
Years of schooling	17.3	1.72	17.1	1.56	17.1	1.55
Years in the USA	16.4	9.98	–	–	–	–
Age at migration	27.0	9.50	–	–	–	–
Annual wage income ($1000)	89.8	95.6	101	95.4	102	99.4
Parental characteristics						
Parental education	–	–	13.8	2.89	14.0	2.69
Parental income ($1000)	–	–	36.7	7.22	37.3	6.55
No. of siblings	–	–	1.21	2.01	1.03	2.04
Ethnic network (%)	.044	.051	.044	.050	.043	.048
Origin characteristics						
Democracy	.572	.451	.720	.416	.662	.452
State of war	.018	.065	.007	.045	.009	.047
% labor in agriculture (%)	40.4	22.3	31.2	25.5	30.1	21.6
GDP per capita ($1000)	4.94	3.88	6.71	4.25	6.90	4.05
Gini coefficient (%)	38.3	15.3	29.2	14.9	27.0	14.9
Mean educational attainment	4.19	2.41	5.38	2.32	5.47	2.22
Distance to the USA (1000 km)	6.35	3.66	6.74	3.60	6.80	2.98
English official language (= 1)	.231	.421	.369	.482	.187	.390
No. of individuals	35,689		18,736		31,199	
No. of origins	49		49		49	

The individual annual wage income is reported in thousands of 2000 US dollars. The parental income is reported in thousands of 1950 US dollars. The origin country's GDP per capita is reported in thousands of 1990 Geary-Khamis dollars. The summary statistics are weighted by the ACS person weight

problem associated with barriers to entering certain narrowly defined jobs. Since each occupation category covers multiple industries, an individual's choice set is less likely to be constrained by the local industrial structure. However, the aggregation inevitably leads to a loss of heterogeneity across occupations.

I inspect five occupational attributes: expected wage income, years of working experience,[14] regular hours of work, training cost, and the likelihood of unemployment.

[14]I measure the years of experience for individuals engaging in various occupations because the census occupational classification includes executive and managerial jobs as a separate category where promotion to these positions largely depends on performance and experience.

These five attributes determine the expected lifetime earnings from working in a job (Boskin 1974). I derive these measures from 1980, 1990, 2000, and 2010 Census data. In each census, I assign a three-digit-coded occupation a value representing the median annual wage income of highly educated individuals who work in that occupation to reflect the expected wage of that occupation.[15] I derive the expected years of experience[16], weekly hours of work, and training costs[17] in the same manner. I calculate the fraction of college graduates who had an occupation but were unemployed at the time of the survey as the likelihood of unemployment. I then average the five measures to the 23 aggregate categories weighted by the number of college graduates working in each narrowly defined occupation.

Ideally, individuals' rational expectation about an occupation is best described by the occupational attributes at the time they choose the job. The start time, however, is not available in the ACS. Because most students graduate from college at the age of 22 to 24 and because many start to work full-time right afterward, I match individuals to the occupational measures derived from a census year that is closest to their potential year of labor market entry.[18]

Table 2 displays the mean attributes of the 23 occupational categories. Among all categories, doctors and lawyers enjoy the highest income and the best job security but require the most training and longest hours of work. General executives, managers, architects, and engineers are also remunerative jobs. As expected, executive and managerial jobs are associated with most years of experience. Household and other service workers, as well as laborers, are paid the lowest. The unemployment rate is the highest among the laborers.

Table 3 presents the fraction of individuals working in each job category by generation. The majority of the work in this sample is in managerial and professional jobs. The share of individuals working as managers and executives (over 20%) is the highest for all three generations. Individuals who are employed in management-related jobs, architect and engineer jobs, and sale representative jobs also make up sizable proportions. The occupational selection patterns are analogous for the 1.5 generation and native-born Americans. Compared to the other two groups, first-generation immigrants are more likely to engage in non-professional jobs or work as professors, scientists, and engineering and science technicians.

[15] The census provides the occupational income score, which assigns each occupation a value representing the median total income of all individuals with that particular occupation. Because I focus on the highly educated individuals, I modify the score to be the median total income score among all college graduates with each occupation.

[16] Work experience is calculated as years of experience = age − years of education − six.

[17] The training cost of each occupation is assigned by the median educational attainment of individuals with a college degree who work in that occupation. Certain occupations, such as physicians and professors, require additional training beyond college.

[18] Individuals who were born between 1956 and 1960 are matched to 1980 occupational attributes; those born between 1961 and 1970 are matched to 1990 attributes; those born between 1971 and 1980 are matched to 2000 attributes; those born after 1980 are matched to 2010 attributes.

Table 2 Occupational attributes in the USA

Occupation	Occupational attributes				
	Wage	Exper.	Hours	Educ.	%U/E
Managerial and professional					
General executives/managers	818	12.8	47.2	16.8	1.47
Management related	667	11.3	44.3	16.7	1.54
Architects/engineers	660	11.2	43.3	16.8	1.19
Professors/scientists	543	10.2	41.8	18.3	1.50
Doctors/lawyers	1060	9.68	50.1	19.5	.590
Health assessors/teachers	488	11.5	42.4	17.1	1.00
Social workers	381	11.2	44.0	17.3	1.61
Writers/artists/athletes	537	10.9	42.4	16.7	3.58
Engineering/science technicians	612	10.7	41.9	16.7	1.86
Health/legal technicians	463	9.42	41.4	17.3	2.10
Non-professional					
Sales representatives	723	11.9	46.4	16.4	1.88
Sales clerks	445	10.6	41.3	16.6	3.87
Office clerks/health service	433	10.2	41.5	16.7	3.65
Administrative support	519	11.8	43.0	16.5	1.78
Protective service	526	11.8	44.4	16.4	1.40
Household/other service	304	10.0	40.0	16.6	5.28
Farmers	379	11.7	50.5	16.4	3.02
Mechanics/system operators	559	12.1	45.1	16.5	2.70
Repairers/precision workers	480	11.7	43.2	16.4	3.32
Construction trades	380	11.3	41.6	16.5	7.10
Heavy machine operators	519	11.7	45.2	16.4	4.46
Machine operators/drivers	389	11.4	42.8	16.5	5.37
Laborers	355	10.5	40.5	16.5	8.74
Mean	532	11.1	43.7	16.9	3.00
Standard deviation	188	1.44	2.87	.740	2.17

Reported occupational attributes are averages across periods. The occupational wage is in 1982–1984 US dollars. The fraction unemployed is displayed in percentage points

5.3 Occupational wage in origin countries

The data on wages by occupation in different countries come from the Occupational Wages around the World Database (OWW) by Richard B. Freeman and Remco H. Oostendorp. This dataset contains wage data for 161 occupations in 171 countries from 1983 to 2008. The average hourly and monthly wage rates for adult workers in each occupation are documented in both local currency units and US dollars.

Table 3 Shares of individuals in various occupational categories

Occupation	First Gen.	1.5 Gen.	Native-born
General executives/managers	21.6	22.9	22.8
Management related	6.32	8.29	8.69
Architects/engineers	11.2	11.39	9.76
Professors/scientists	8.15	4.92	4.76
Doctors/lawyers	5.10	8.53	8.60
Health assessors/teachers	4.02	6.35	7.29
Social workers	1.95	2.31	2.12
Writers/artists/athletes	2.68	3.20	3.54
Engineering/science technicians	7.18	6.07	4.98
Health/legal technicians	1.12	1.03	.830
Sales representatives	7.02	8.30	9.56
Sales clerks	3.19	2.76	2.80
Office clerks/health service	2.18	2.04	2.05
Administrative support	1.36	2.11	2.01
Protective service	1.32	2.74	2.64
Household/other service	3.59	1.25	1.32
Farmers	.790	0.55	1.05
Mechanics/system operators	2.98	1.86	2.02
Repairers/precision workers	.890	.530	.490
Construction trades	1.75	.590	.660
Heavy machine operators	.550	.400	.380
Machine operators/drivers	3.70	1.44	1.32
Laborers	1.35	.610	.530
No. of individuals	35,689	18,736	31,199
No. of origins	49	49	49

I employ hourly wages with country-specific calibration in US dollars. To ensure that occupational wages are comparable across nations and over time, I adjust the data using the purchasing power parity conversion factor obtained from the World Bank International Comparison Program Database. Since the absolute wage rate may capture the cross-country income disparity, I normalize the wage rates to have a mean of zero and a standard deviation of one within each country per year.

Next, I convert the classification of occupations in the OWW data to the 1990 Census Bureau occupational classification and group the narrowly defined occupations into the 23 categories discussed previously. I use the mean wage rate in a category as a measure of the categorical wage rate. I link each birth cohort in the ACS sample to the country-of-origin wage distribution during the period when the cohort members

were about 25 years old, around the time that they entered the labor force.[19] Using past country-of-origin wages may also help minimize the effect of potential return migration on current occupational choices.

Table 4 shows the summary statistics of occupational wages in the source countries covered by the ACS sample, with and without standardization. Forty-nine countries are included. There is substantial heterogeneity in both the actual and standardized categorical wage rates. While the majority of heterogeneity is between countries, the variation over time is also sizable. Figure 1 depicts the average standardized wages between 1983 and 2005 in eight countries that have the most observations in the ACS sample; the figure shows a considerable discrepancy in the wage distribution across nations. The occupational wage distributions in the 49 countries deviate from the US wage distribution to various degrees. The correlation between US wages and standardized wages in other countries has a mean of 0.6 with a minimum of 0.09 (Costa Rica, 1983–1985) and a maximum of 0.85 (Finland, 1986–1990).

Admittedly, ethnic groups in the USA may face differential compensating wages. If the origin-specific wage distribution in the US market is correlated with the wage distribution in one's country of origin, the latter may pick up the effect of the former. To address this issue, I calculate the US occupational wages by ethnic origin in the same manner as I calculate the national average US wage. Figure 2 exhibits the correlation between the origin-specific US wage and the national average US wage in teal and the correlation between the origin-specific US wage and the wage in the origin country in red. The origin-specific wage distributions in the USA are more comparable to the national average US wage distribution than are those in other countries.

5.4 Other origin characteristics

To better rule out the probability that the origin wage income picks up the effect of unobserved heterogeneity across ethnic origins, I collect a number of origin-country characteristics that may affect selective migration. These characteristics include democracy status, state of war, share of the labor force in agriculture, GDP per capita, Gini coefficient, average educational attainment among the adult population, distance to the USA, and whether English is an official language (Betts and Lofstrom 2000; Borjas 1987; Lewer and den Berg 2008).

The data on these characteristics are gathered from various sources.[20] Except for distance to the USA and whether English is an official language, I calculate the averages for these variables for each period specified for the OWW data and match the averages to the six birth cohorts in the ACS sample. Summary statistics of these origin characteristics are displayed in Table 1.

[19] Explicitly, I link birth cohort 1956–1960 to origin wages (averaged by period) in 1983–1985, cohort 1961–1965 to origin wages in 1986–1990, cohort 1966–1970 to origin wages in 1991–1995, cohort 1971–1975 to origin wages in 1996–2000, cohort 1976–1980 to origin wages in 2001–2005, and cohort 1981–1985 to origin wages in 2006–2008, respectively.

[20] More details about the data sources are provided in the Appendix Section A.3.

Table 4 Summary statistics of occupational wages in origin countries

	Actual wage				Standardized wage			
		Standard deviation				Standard deviation		
	Mean	All	Between	Within	Mean	All	Between	Within
Overall	4.00	4.44	–	–	.342	1.03	–	–
Occupation								
General executives/managers	6.25	5.77	4.89	2.67	1.69	1.20	.805	.914
Management related	4.50	4.92	4.17	2.17	.804	.837	.530	.642
Architects/engineers	5.47	4.93	4.41	1.65	1.43	.807	.700	.543
Professors/scientists	6.66	6.55	5.45	2.96	1.78	1.27	.941	.807
Doctors/lawyers	7.07	6.96	5.99	2.59	1.71	1.19	.947	.679
Health assessors/teachers	4.47	4.34	3.86	1.46	.514	.558	.473	.351
Social workers	3.78	3.60	3.14	1.37	.096	.603	.484	.357
Writers/artists/athletes	5.02	4.69	4.09	1.63	.791	.654	.510	.457
Engineering/science technicians	5.80	5.44	4.62	2.35	1.47	1.13	.757	.823
Health/legal technicians	3.47	3.70	3.18	1.44	−.043	.543	.440	.354
Sales representatives	3.13	3.34	2.89	1.29	−.195	.457	.345	.299
Sales clerks	2.95	3.16	2.76	1.12	−.218	.401	.387	.235
Office clerks/health service	3.14	3.32	2.95	1.09	.047	.504	.368	.373
Administrative support	3.40	3.53	3.10	1.31	−.016	.393	.285	.260
Protective service	4.38	4.24	3.59	1.64	.423	.746	.546	.500
Household/other service	2.35	2.60	2.38	.725	−.647	.304	.259	.167
Farmers	2.45	2.49	2.19	1.06	−.512	.444	.398	.281

Table 4 (continued)

	Actual wage				Standardized wage			
		Standard deviation				Standard deviation		
	Mean	All	Between	Within	Mean	All	Between	Within
Mechanics/system operators	3.23	3.46	3.16	1.05	−.037	.250	.204	.159
Repairers/precision workers	3.22	3.31	2.93	1.10	0.049	.666	.496	.400
Construction trades	2.88	3.31	3.04	1.00	−.317	.298	.256	.190
Heavy machine operators	3.22	3.59	3.34	1.00	−.072	.365	.343	.219
Machine operators/drivers	2.75	3.12	2.88	.896	−.355	.218	.186	.163
Laborers	2.49	3.03	2.81	.890	−.482	.279	.206	.212
No. of observations				4664				4664
No. of origins				49				49
No. of periods				6				6

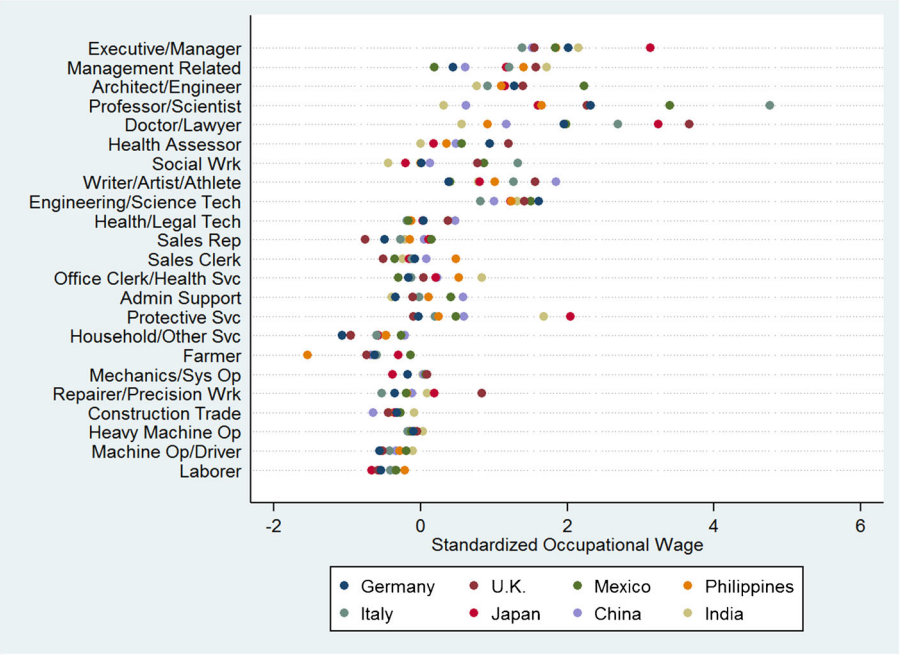

Fig. 1 Occupational wage distributions in eight countries—standardized wage

6 Origin wage distributions and occupational choices

6.1 Main results

This section investigates the link between the occupational wage distribution in origin countries and the job selection of individuals in the USA. For easier comparison, in all the regressions that follow, I normalize the five occupational attributes measured in the USA to have a mean of zero and a standard deviation of one across categories and cohorts; I normalize the standardized origin occupational wage to have a mean of zero and a standard deviation of one across countries, periods, and categories.

I begin with a parsimonious specification that includes only the five US occupational attributes and then introduce the origin occupational wage to the regression. I perform the same set of regressions on the three generations and report the results in Table 5. All the regressions are estimated by the conditional logit model and weighted by the ACS person weight. The robust standard errors are clustered at the origin-cohort level.

As shown in the odd-numbered columns in Table 5, the relative weights placed on the five attributes by different generations are analogous and consistent with the findings in the literature (Boskin 1974; Dolton et al. 1989; Arcidiacono 2004). Specifically, individuals view monetary income as positive but training costs and

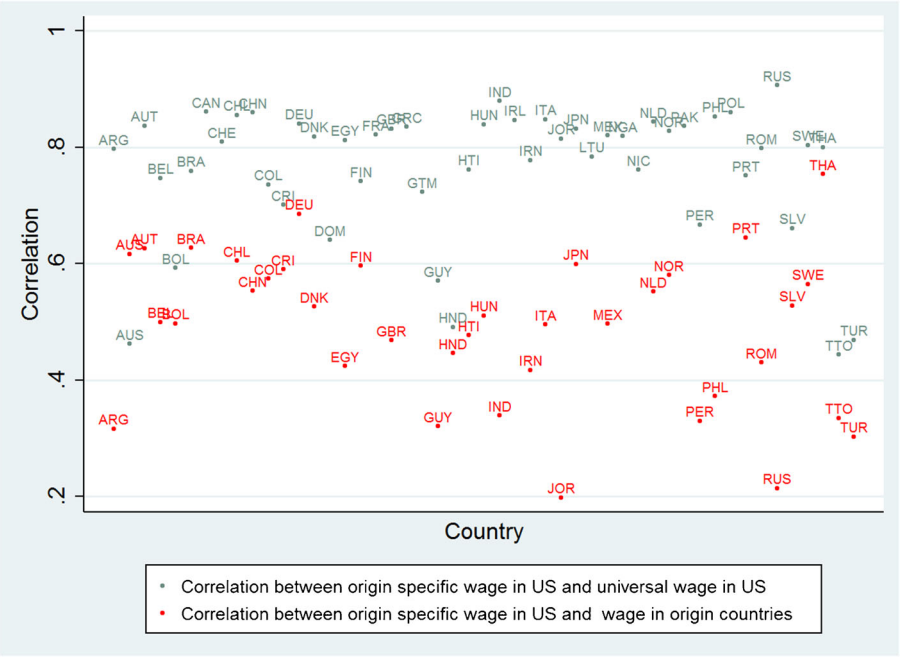

Fig. 2 Correlations between occupational wages in the USA and origin countries

hours of work as negative. Compared to later generations, first-generation immigrants disvalue required years of education and the chance of unemployment less, presumably as a result of the obstacles they encounter in the labor market. This pattern may align with the phenomenon of immigrants' downgrading discussed in Dustmann and Preston (2012) and Dustmann et al. 2016).

The origin occupational wage displays a positive and significant impact on the job selection of all three generations. The estimates suggest that individuals are more likely to go for a job which is associated with higher monetary rewards in their origin society, conditional on the job characteristics they directly confront in the USA. The importance of the origin wage declines over generations. The estimated coefficient is significantly larger for the first generation than the two other generations.

Since the magnitudes of the estimates from the conditional logit model are not readily interpretable, I calculate the change in the predicted probability to select one of the 23 occupation categories, if, *ceteris paribus,* the wage of that category in the country of origin grows by one standard deviation. The simulated data suggest that a one standard deviation increase in the origin wage predicts the average likelihood to choose a job category to grow by 0.5 percentage point (11%) for a first-generation immigrant, 0.2 percentage points (5%) for a 1.5-generation immigrant, and 0.1 percentage points (3%) for a native-born American. The predicted marginal effects might verify that origin occupational wage distribution plays an economically

Table 5 Origin wage distribution and occupational choices: baseline

	First generation		1.5 Generation		Native-born	
Variable	(1)	(2)	(3)	(4)	(5)	(6)
Origin occupational wage		.110***		.052***		.034***
		(.014)		(.015)		(.012)
US occupational wage	.902***	.781***	.994***	.933***	.896***	.861***
	(.065)	(.059)	(.057)	(.060)	(.035)	(.036)
Years of experience	.214***	.179***	.244***	.224***	.246***	.236***
	(.034)	(.035)	(.034)	(.036)	(.018)	(.018)
Hours of working	−.531***	−.426***	−.511***	−.455***	−.391***	−.359***
	(.057)	(.054)	(.037)	(.043)	(.025)	(.025)
Education required	−.122***	−.153***	−.217***	−.240***	−.206***	−.218***
	(.035)	(.034)	(.043)	(.043)	(.023)	(.024)
% unemployed	−.024	.002	−.492***	−.483***	−.506***	−.500***
	(.032)	(.032)	(.033)	(.032)	(.020)	(.020)
Marginal effect		.470		.212		.138
		[11.0%]		[5.09%]		[3.32%]

Table 5 (continued)

	First generation		1.5 Generation		Native-born	
Variable	(1)	(2)	(3)	(4)	(5)	(6)
Difference in $\widehat{\phi}_1$ between generations						
$t_{\Delta 1,1.5}$						4.04***
$t_{\Delta 1,N}$						5.79***
$t_{\Delta 1.5,N}$						1.36
No. of observations	818,372	818,372	429,915	429,915	715,817	715,817
No. of origins	49	49	49	49	49	49
Log-likelihood/10^6	−11.8	−11.8	−5.21	−5.21	−8.56	−8.56

*Significant at 10%; **Aignificant at 5%; ***Significant at 1%. Regressions are estimated using a conditional logit model, and observations are weighted by the ACS person weight. The dependent variable is an indicator of choice among the 23 occupation categories. Robust standard errors in parentheses are clustered at origin-cohort level. The reported marginal effects are the average percentage point change in choosing one occupational category given a one SD increase in the standardized occupational wage of that category in the origin country. The marginal effects are in percentage points, with percentage changes reported in brackets

important role in the occupational choices of immigrants and their offspring in the receiving country.

Next, I allow the weights placed on occupational attributes to vary by person so that the heterogeneous motivation or endowments across individuals can be taken into account. For straightforward interpretation, I normalize the individual-specific characteristics (except for the year and country-of-origin dummies) to have a mean of zero and a standard deviation of one by generation. Hence, the main effects of the occupational attributes, including origin occupational wage, stand for the weights that an individual of mean characteristics in a generation (from the baseline country, i.e., UK) assigns to them in the baseline year (2000).

I start by controlling for an individual's age, disability, English proficiency, years in the USA (for the first generation only), and year of survey. The interactions between occupational attributes and parental characteristics (i.e., parental income, education, and number of siblings),[21] origin country characteristics (i.e., the GDP per capita, Gini coefficient, average educational attainment, war, democracy, share of the labor force in agriculture, distance to the USA, and whether English is an official language), and origin-country dummies are added to the regressions one by one. I incorporate the parental ethnic network in the latter three specifications as well. Table 6 presents the estimated main effects of the USA and country-of-origin wages. The estimates on the origin occupational wage, as well as the simulated marginal effects in columns 1–3, resemble those in Table 5. However, when country-of-origin fixed effects are introduced in the last column, the standard errors increase dramatically, and the estimated coefficients become insignificant, possibly due to a lack of temporal variation in wage distributions of origin countries.

As discussed in Section 3 (Conceptual Framework), a positive link between the origin wage and one's US job choice may result from unobserved advantages (due to specialization of higher-paying jobs or the positive selection of migrants), the prospect of return migration, cultural norms (if a high-earning job is associated with a high cultural status in the source country), or information frictions. Notably, both the 1.5 generation and native-born Americans probably have no first-hand experience of their country-of-origin labor markets. Nor would many of them view the origin countries as future destination markets. Therefore, conditional on family backgrounds, ethnic networks, and selective migration, origin occupational wages are less likely to affect the two later generations through origin-specific human capital or potential return migration but are more likely to exert an effect via norms or misinformation.

The interpretation, however, is more complicated for the first generation. First, there is substantial heterogeneity in the education path of first-generation immigrants that the current control set may fail to capture. Second, the job selection of non-US citizens may be restricted by US immigration policies.[22] Occupations of first-generation immigrants may reflect the preferences or ability of US employers to hire

[21] The parental characteristics are not in the control set for the first-generation immigrants as the group means derived from the population in the USA are potentially poor proxies.

[22] In the sample, about 45% of first-generation immigrants have naturalized, and about 87% of 1.5 generation have naturalized.

workers from certain countries via H-1B visas. As a result, occupational choices of the first generation may not be comparable to those of the other two generations.[23]

6.2 Robustness

I explore modifying the regressions in Tables 5 and 6 in various ways to test the robustness of the results. First, to better disentangle the intergenerational human capital transmission from other channels, I use an alternative dataset, the 1972–2008 General Social Survey (GSS). The GSS is a much smaller dataset but allows the controls of respondents' parental education, occupation, and other family characteristics. Second, I employ a nested logit model that does not impose the independence of irrelevant alternatives (IIA) assumption as the conditional logit model. The former also addresses the concern that estimates from the conditional logit model primarily pick up the difference across broad nests instead of across occupations. Third, to eliminate the possibility that origin wages capture measurement errors in US wages, I conduct two tests: (1) I perform a falsification test that randomly assigns "origin" and "cohort" to individuals and links them to the origin wages based on the given origin and cohort; (2) I replace US occupational wages constructed from the census data by US wages from the OWW database. The latter may be more analogous to the origin wages derived from the same data source. Finally, to justify the assumption that occupational wages in the origin country at the age of 25 affect one's current occupation, I test specifications where I employ the origin wage of alternative periods: the wages during the period of emigration for the first generation, wages prior to college, and the current wages that may be more relevant than previous wages to potential return migrants. All the robustness checks yield reassuring results.[24]

6.3 Assimilation

Immigrants assimilate into a receiving society with the time they spend there. All four channels discussed previously may predict a declining part of the origin occupational wage in immigrants' US job choices. A slow acquisition of US-specific human capital and information about local labor markets, as well as a growing identification with US culture, may help increase the relevance of immigrants' skills in their occupational choice and boost the economic returns to such skills. Return migration may also become less likely over time as immigrants become more attached to the USA, both economically and socially.

Table 7 inspects the role of immigrant assimilation, as captured by the number of years a first-generation immigrant has lived in the USA. The first two columns repeat

[23] In response to the two abovementioned issues, I replicate the regressions in columns 2–3 Table 6 (1) excluding individuals who migrated before the age of 18, as there may be substantial heterogeneity in education and motivation for migration among this group and (2) restricting the sample to the first-generation immigrants who have naturalized. Appendix Table 12 presents the results. The estimates are comparable to those in Table 6 and may thus suggest that neither differential education path nor visa policies drive the positive link between origin wages and first generation's occupational choices in the USA.

[24] More details about these tests are available in the online appendix.

the regressions in columns 2–3 in Table 6.[25] The last two columns replace the number of years in the USA with the log number of years. As expected, the interaction effect of origin occupational wage is negative and significant in all the specifications. When individual characteristics, ethnic networks, and origin characteristics are accounted for, column 2 suggests a 1% decrease in the importance of origin wage if an immigrant stays in the USA for an additional year; column 4 suggests a 10% decrease given a one standard deviation increase in length of time in the USA from the sample mean (a 61% increase). These modest changes may imply slow economic convergence of immigrants, as argued by Abramitzky et al. (2014, 2020).

Notably, departures of negatively selected return migrants may be another explanation besides the previously referenced channels (Abramitzky et al. 2020). A higher weight placed on the origin occupational wage may imply a more serious skill-job mismatch and thus a lower income. As a result, immigrants who rely more on their country-of-origin wage distribution may display a higher chance of return migration and a shorter stay in the USA.[26]

7 Potential mechanisms

This section explores the possible channels through which wage distributions in origin countries affect the occupational choices of individuals in the USA. Given data limitations, the current analysis can differentiate among a few of the potential mechanisms, but not among all of them. I will discuss this issue in more depth in Section 9.

7.1 Unobserved advantage

In considering the positive link between origin wages and immigrants' occupational choices in the USA, one possibility is that the country-of-origin wage standing represents the unobserved opportunity or advantage of a group to work in an occupation that the current control set fails to capture. Therefore, in this section, I investigate whether the origin wages are predictive of a person's actual salary conditional on the average wage structure of US occupations. If individuals appear to earn more in an occupation that pays relatively better in their origin country, the positive effect of origin wages may be explained by inherited traits, job specialization, selective migration, or networks that make one ethnic group more productive than other groups in that occupation.

[25] The regressions in columns 2–3 Table 6 control for the years in the USA by including the interactions between the six occupational attributes and the standardized number of years in the USA. For a more straightforward interpretation, I use the actual number of years in the USA in this table. So the estimated main effects and interaction effects of the origin wage are slightly different.

[26] In a similar vein, I investigate if the effect of the origin wage varies by age at migration. I find that the link between origin wage and US occupational choice is stronger if individuals migrated at a later age. The results are available in the online appendix.

Table 8 presents regression results. The dependent variable is a person's annual wage and salary income. The variable of interest is the wage of an individuals' chosen occupation in the country of origin. The US wage of that occupation, as well as the same set of individual characteristics, parental characteristics, and origin characteristics (see Table 6), is controlled for. The even-numbered columns include the origin fixed effects, but the odd-numbered columns do not. All the regressions are estimated by OLS, weighted by the ACS person weight, and the standard errors are clustered at the origin-cohort level.

Table 6 Origin wage distribution and occupational choices: additional controls

	(1)	(2)	(3)	(4)
First generation				
Origin occupational wage	.123***	.119***	.131***	.111
	(.014)	(.015)	(.013)	(.092)
US occupational wage	.956***	.918***	.957***	.631
	(.068)	(.063)	(.056)	(.424)
Marginal effect	.525	.498	.547	.455
	[12.4%]	[12.0%]	[13.3%]	[11.1%]
No. of observations	818,372	818,372	818,372	818,372
Log-likelihood/10^6	−11.5	−11.2	−11.1	−10.9
1.5 generation				
Origin occupational wage	.043**	.050***	.085***	.147
	(.021)	(.019)	(.017)	(.103)
US occupational wage	1.08***	.966***	.959***	−.040
	(.072)	(.062)	(.051)	(.381)
Marginal effect	.176	.203	.345	.614
	[4.23%]	[4.94%]	[8.42%]	[15.1%]
No. of observations	429,915	429,915	429,915	429,915
Log-likelihood/10^6	−5.19	−5.13	−5.11	−5.06
Native-born				
Origin occupational wage	.027	.041**	.053 ***	−.057
	(.018)	(.016)	(.017)	(.139)
US occupational wage	.955***	.843***	.833***	−.275
	(.056)	(.050)	(.048)	(.332)
Marginal effect	.109	.165	.213	−.218
	[2.62%]	[4.00%]	[5.16%]	[5.28%]
No. of observations	715,817	715,817	715,817	715,817
Log-likelihood/10^6	−8.53	−8.46	−8.43	−8.37
Parental characteristics	No	Yes	Yes	Yes
Ethnic networks	No	Yes	Yes	Yes
Origin characteristics	No	No	Yes	Yes
Origin fixed effects	No	No	No	Yes

Table 6 (continued)

	(1)	(2)	(3)	(4)
Difference in $\widehat{\phi}_1$ *between generations*				
$t_{\Delta 1,1.5}$	4.78***	4.18***	3.17***	.375
$t_{\Delta 1,N}$	5.45***	5.04***	5.20 ***	1.44
$t_{\Delta 1.5,N}$	.834	.524	1.88*	1.61

*Significant at 10%; **Significant at 5%; ***Significant at 1%. Regressions are estimated using a conditional logit model, and observations are weighted by the ACS person weight. The dependent variable is an indicator of choice among the 23 occupation categories. Robust standard errors in parentheses are clustered at origin-cohort level. Individual characteristics are controlled for in all the specifications. Non-alternative-specific variables are introduced to the regressions by interacting them with the six occupational attributes. For first-generation immigrants, parental income, education, and the number of siblings are not included in the control set. The reported marginal effects are the average percentage point change in choosing one occupational category given a one SD increase in the standardized occupational wage of that category in the origin country. The marginal effects are in percentage points, with percentage changes reported in brackets

Table 7 Heterogeneous effects by years in the USA

	First generation			
Variable	(1)	(2)	(3)	(4)
Origin occupational wage	.155***	.154***	.244***	.226***
	(.023)	(.020)	(.031)	(.029)
Origin wage×years in the USA	−.002***	−.001*		
	(.001)	(.001)		
Origin wage×ln(years in the USA)			−.049***	−.036***
			(.010)	(.009)
US occupational wage	.918***	.957***	.921***	.958***
	(.063)	(.056)	(.063)	(.056)
No. of observations	818,372	818,372	818,372	818,372
Log-likelihood/10^6	−11.2	−11.1	−11.2	−11.1
Individual characteristics	Yes	Yes	Yes	Yes
Ethnic networks	Yes	Yes	Yes	Yes
Origin characteristics	No	Yes	No	Yes
Origin fixed effects	No	No	No	No

*Significant at 10%; **Significant at 5%; ***Significant at 1%. Regressions are estimated by the conditional logit model, weighted by the ACS person weight. The first-generation immigrants are examined. The dependent variable is an indicator of choice among the 23 occupation categories. Robust standard errors in parentheses are clustered at origin level. Non-alternative-specific variables are introduced to the regressions by interacting them with the six occupational attributes

Table 8 Individual wage income

	First generation		1.5 generation		Native-born	
Variable	(1)	(2)	(3)	(4)	(5)	(6)
Origin occupational wage	.702	.554	−1.50	−1.57	−2.44***	−1.92***
	(.589)	(.567)	(1.40)	(1.46)	(.904)	(.678)
US occupational wage	24.2***	23.3 ***	29.8***	29.3***	29.8***	28.9***
	(1.11)	(1.16)	(2.20)	(2.29)	(1.14)	(1.05)
Age	.863***	.505**	.790**	.820*	.518	1.40 ***
	(.208)	(.209)	(.325)	(.424)	(.315)	(.291)
1 (disability)	−14.5***	−13.5***	−15.8***	−16.4***	−22.7 ***	−23.5***
	(3.64)	(3.43)	(3.07)	(3.20)	(2.70)	(2.70)
1 (speaking English)	14.4***	11.8 ***	14.1***	14.3***	12.6**	11.9**
	(1.76)	(1.46)	(4.86)	(4.81)	(5.51)	(5.91)
Years in the US	−.008	.106				
	(.134)	(.144)				
Parental education			.027	−1.39	−.015	−2.43*
			(.772)	(1.517)	(.825)	(1.26)
Parental income			.001	.001*	.001**	.000
			(.000)	(.001)	(.000)	(.000)
No. of siblings			−.692*	−.666 *	−.249	−.267
			(.377)	(.388)	(.304)	(.271)
Ethnic network (%)	38.5**	25.2*	−15.48	−23.5	−20.4	−23.6
	(14.6)	(13.1)	(33.3)	(29.8)	(18.0)	(15.6)
Origin fixed effects	No	Yes	No	Yes	No	Yes
No. of individuals	26,656	26,656	14,058	14,058	23,341	23,341
R-squared	.212	.228	.195	.202	.177	.187

*Significant at 10%; **Significant at 5%; ***Significant at 1%. Regressions are estimated using OLS, and observations are weighted by the ACS person weight. The dependent variable is personal annual wage and salary income (measured in thousands of 2000 US dollars). Robust standard errors in parentheses are clustered at origin level. The regressions also control for year fixed effects, state fixed effects, and origin-country characteristics. The even-numbered columns include origin fixed effects, but the odd-numbered columns do not

The estimated coefficient on the origin wage is positive but insignificant for the first generation. Notably, the estimates for the first generation may be positively biased because the controls for parental characteristics are not present. Among the later generations, however, engaging in an occupation that was more profitable in the origin country predicts lower actual earnings in the USA, though the estimate is only statistically significant for the native-born. In particular, one standard deviation increase in the origin occupational wage is associated with about $1,900 to $2,500 less in a native-born American's annual income, representing 2–3% of the sample average wage and salary income. These results may help rule out the probability that

the unobserved advantage explains why individuals in the USA self-select into jobs that were more lucrative in their countries of origin.

There may be two explanations for the negative connection between origin wage and individuals' realized US income. First, consistent with the theory that people migrate to maximize the returns to their skills (Borjas 1987), adverse selection may exist in occupations that were better-paid in the countries of origin, whereas positive selection may occur for the lower-paid occupations. Second, keeping return migration as an option, immigrants may have a lower incentive to assimilate and to invest in human capital in the USA (Schaeffer 1995). The more lucrative the current occupation is back home, the better the outside option that individuals face, and the lower incentive they have to accumulate human capital. Deficient human capital investment leads to lower productivity.

7.2 Knowledge and information

This section uses the Knowledge of World of Work (KWW) survey to probe the role of information. The KWW is a part of the National Longitudinal Survey of Youth 1979 Children and Young Adults (NLSY79 Child/YA). The survey asked individuals aged 14–24 a series of questions concerning commonly held jobs. Presumably, these people had limited hands-on experience with these jobs at the time of the survey. Their impression of occupations may depend on how they were taught by their parents, as well as the influence of their ethnic community. If the ways in which young individuals view different occupations is associated with the occupational wages in their countries of origin, it may be information transmission that explains the connection between origin wage distribution and job selections of immigrants and their descendants in the USA.

I focus on a set of six questions which ask the respondents to pick one out of two occupations that they think offers a higher wage. The six questions cover nine of the 23 occupational categories.[27,28] Because the KWW questions compare occupations by pair, ranking the perceived earnings of the nine occupational categories is difficult. Nevertheless, within each pair, if an ethnic group considers option A pays more than option B, the share of the young people in that group choosing A over B would be larger. The greater the perceived earning gap between these two jobs, the more significant the difference in the propensities to select one over the other.

I pool the answers to the six KWW questions together. Since choosing one option in a question necessarily means not choosing the other, I start by focusing only on

[27] Mentioned occupational categories include the following: (1) Architects/engineers, (2) doctors/lawyers, (3) health assessors/teachers, (4) sales clerks, (5) protective service, (6) household/other service, (7) mechanics/system operators, (8) construction trades, and (9) machine operators/drivers. More detailed information about the survey is provided in Appendix Section A.4.

[28] As before, I dismiss the origins with fewer than 10 observations. A total of 2215 individuals from 12 countries of origin are included accordingly, both males and females. Appendix Table 11 displays the included origin countries. Despite the much smaller set of origin countries, these countries represent the ones with a large number of observations in the ACS. Moreover, members of the KWW sample were born between 1970 and 1983 and may resemble the fourth and fifth cohorts in the ACS sample. But I cannot tell whether they were foreign- or native-born.

whether a respondent selects option A (as listed in the KWW questions). I generate a binary indicator and regress it on the US and country-of-origin wages[29] of the occupation in question (i.e., option A) and those of the occupation of comparison (i.e., option B). The results are shown in columns 1–3 of Table 9. Columns 1 and 2 control for respondents' gender, age, and the year of survey; column 3 controls for individual fixed effects. I also add to the control set the question fixed effects in the latter two columns. The regressions are estimated by OLS, and the robust standard errors are clustered at the origin-by-question level. The answers to the KWW questions conform to the wage distribution in the USA: If the US wage is held constant, the origin wage appears predictive of how individuals rank occupational earnings in the KWW. In particular, a one standard deviation increase in the origin wage of option A leads to a 0.7 percentage point increase in the probability that a respondent chooses A over B. The estimates on the origin wage of option B are not significant.

In case the designation of "option A" listed in a question affects the estimation results, I consider whether individuals pick one specific occupation (either A or B) as the outcome. I then regress the choice on US and country-of-origin wages of both occupations in the question. I replicate the regressions in the first three columns and display the results in columns 4–6 of Table 9. The results suggest that respondents would view an occupation as more lucrative in the USA if the occupation was associated with a higher country-of-origin wage rate, or if the comparison was less financially rewarding in the origin country: given a one standard deviation increase in the origin wage of an occupation (relative to the comparison occupation), the chance that an individual views the occupation as more profitable than the comparison increases by 16 percentage points, a change much larger than that estimated in the first three columns. These results may support the argument of information friction.[30]

Nevertheless, two caveats are worth noting. First, the sample surveyed in the KWW is rather small. Only a limited number of origin countries and occupations are involved. Second, the KWW questions ask respondents to compare the earnings of narrowly-defined jobs rather than occupational categories. Though the wage income of jobs within the same category may be relatively comparable, the jobs referenced in the questions do not necessarily offer a wage close to the categorical mean.

8 Individuals with different educational levels

Individuals' educational decisions can be endogenous to their occupational choice. It is possible that the college attendance rate of the US population varies by national

[29] I use the average standardized occupational wages from 1983 to 1995 in the regressions in Table 9.

[30] I conduct two sensitivity tests to rule out the possibility that the origin occupational wages capture additional relevant information that is not incorporated in the average US wage, such as regional wage differences. First, I replace the national averages of US wages using the regional averages in the analysis just described. Four US regions are considered as follows: Northeast, North Central, South, and West. Second, I restrict attention to occupation pairs with unambiguous wage ranking across states in the USA and exclude the answers to questions 1 and 3, accordingly. These tests provide reassuring results which are available in the online appendix.

Table 9 KWW answers and origin occupational wages

	1 (choosing option A)			1 (choosing certain occupation)		
Variable	(1)	(2)	(3)	(4)	(5)	(6)
Occupation in question						
Origin occupational wage	.266***	.008**	.007*	.162***	.162***	.162***
	(.076)	(.003)	(.004)	(.051)	(.051)	(.053)
US occupational wage	.034	.771***	.768***	.072*	.071*	.071*
	(.056)	(.109)	(.114)	(.036)	(.036)	(.038)
Occupation of comparison						
Origin occupational wage	−.008	-.000	−.003	−.162***	−.162***	−.162***
	(.108)	(.004)	(.004)	(.051)	(.051)	(.053)
US occupational wage	−2.45***	−2.15 ***	−2.15***	−.073*	−.074**	−.074*
	(.090)	(.055)	(.064)	(.036)	(.037)	(.038)
Question fixed effects	No	Yes	Yes	No	Yes	Yes
Individual fixed effects	No	No	Yes	No	No	Yes
F(Origin wages of both jobs = 0)	6.41 ***	2.67*	2.61*	5.43***	6.05***	4.96***
No. of observations	13,290	13,290	13,290	26,580	26,580	26,580
No. of individuals	2215	2215	2215	2215	2215	2215
No. of origins	12	12	12	12	12	12
R-squared	.327	.619	.681	.283	.283	.284

*Significant at 10%; **Significant at 5%; ***Significant at 1%. Regressions are estimated using OLS. The dependent variable is a binary indicator which is one if a respondent selects a specific occupation (option A in columns 1–3) in a question and zero otherwise. Robust standard errors in parentheses are clustered at origin-by-KWW-question level. Individual's age, gender, and the year of survey are controlled for in columns 1, 2, 4, and 5

origin in a way that is relevant to country-of-origin wage distribution. Also, Dustmann and Preston (2012) point out that immigrants tend to downgrade in their occupations upon arrival, and this phenomenon is particularly stronger for skilled workers. Hence, restricting the sample to individuals who hold a college degree or above may impose interpretation issues.

The less-educated make up an essential portion of immigrants. The relationship between origin-country wages and US occupational choice may vary by education. For instance, less-educated immigrants may encounter greater language barriers and have more difficulties learning about the local labor market than better-educated ones do. Information frictions can therefore be a more severe problem for low-skilled immigrants to find employment. Thus, they may rely on their origin wages more in selecting jobs. Meanwhile, immigrants with a lower level of human capital are more likely to change occupation upon arrival based on the existing opportunities of their ethnic networks (Lafortune and Tessada 2012). If so, origin occupational wage may appear less important to these people.

Accordingly, I expand the sample to include individuals with at least a high school degree or a certificate of high school equivalency (GED) in this section. I choose a high school degree as the sample criteria for two reasons. First, the job choices available to high school dropouts can be minimal. Their actual jobs are more likely a result of access instead of a discretionary decision. Second, examining people with at least a high school degree sets a lower boundary on the unobserved ability so that comparison within the sample is meaningful.

The set of possible occupations available to individuals can be confined by their know-how, market experience, location, and social networks. Many of these factors are unobservable to econometricians. Compared to college graduates, low-skilled workers may face a more restricted choice set because of their qualifications and ability.[31] However, since each occupational category in the data covers multiple narrowly defined jobs, most categories involve workers of all educational levels. So excluding certain occupations from the choice set for workers with less education would be inappropriate.

Consequently, I employ a nested logit model instead of a conditional logit model in this section. A nested logit model assumes choice proceeds through a set of "nested," similar alternatives rather than regarding all occupations as elements of a single choice set. The 23 job categories are grouped into two nests: professional jobs and non-professional jobs. Professional jobs are generally linked to higher income and more years of training than non-professional jobs. I assume years of schooling and other individual characteristics, including age, disability, English proficiency, and years in the USA (for the first generation only), determine the propensity to choose a nest, while origin occupational wage matters to choice within a nest conditional on the same set of parental characteristics, ethnic network, and origin characteristics as

[31] The Bureau of Labor Statistics (BLS) provides information about education and training requirements for narrowly defined occupations. Based on the projections data of BLS on occupational entry-level education, over three-quarters of professional jobs require a bachelor's degree or above, while less than 2% of non-professional jobs require a bachelor's degree. Source: https://www.bls.gov/emp/tables/education-and-training-by-occupation.htm

Table 10 High school graduates and above: nested logit

	First generation		1.5 generation		Native-born	
	Occ. choice	Prof. nest	Occ. choice	Prof. nest	Occ. choice	Prof. nest
Variable	(1)	(2)	(3)	(4)	(5)	(6)
Origin occupational wage	.297***		.305***		.093***	
	(.017)		(.038)		(.026)	
US occupational wage	.961***		1.96 ***		1.15***	
	(.041)		(.103)		(.059)	
Years of experience	−.747***		−.436***		−.565***	
	(.025)		(.044)		(.031)	
Hours of working	.238***		.032		.485***	
	(.036)		(.084)		(.051)	
Education required	−1.08***		−.788 ***		−.652***	
	(.031)		(.065)		(.040)	
% unemployed	−.586***		−1.51 ***		−1.03***	
	(.041)		(.103)		(.066)	
Age		.280***		.284***		.231***
		(.033)		(.063)		(.039)
1 (disability)		.025		.180***		.077**
		(.029)		(.059)		(.037)
1 (speaking English)		.349***		.419***		.061*
		(.033)		(.059)		(.036)

Table 10 (continued)

	First generation		1.5 generation		Native-born	
	Occ. choice	Prof. nest	Occ. choice	Prof. nest	Occ. choice	Prof. nest
Variable	(1)	(2)	(3)	(4)	(5)	(6)
Years of schooling		.422***		.330***		.369***
		(.010)		(.012)		(.010)
Years in the USA		−.245***				
		(.032)				
Marginal effect		.475		.290		.122
		[11.3%]		[6.70%]		[2.79%]
Difference in $\widehat{\phi}_1$ *between generations*						
$t_{\Delta 1,1.5}$						.292
$t_{\Delta 1,N}$						9.43***
$t_{\Delta 1.5,N}$						6.67***
No. of observations	1,287,977		861,166		1,112,602	
No. of individuals	59,210		39,458		52,006	
Log-likelihood/10^4	-15.7		-10.6		-13.8	

*Significant at 10%; **Significant at 5%; ***Significant at 1%. Regressions are estimated using a nested logit model, and observations are weighted by the ACS person weight. Robust standard errors in parentheses are clustered at origin level. The choice of the nest (professional vs. non-professional) is assumed to depend on individual characteristics. Origin occupational wage, parental, and origin characteristics are controlled for in all the specifications and are assumed to affect occupational choices within a nest. For first-generation immigrants, parental income, education, and the number of siblings are not included in the control set. All the US job characteristics are derived by education level. The reported marginal effects are the average percentage point change in choosing one occupational category given a one SD increase in the standardized occupational wage of that category in the origin country. The marginal effects are in percentage points, with percentage changes reported in brackets

those controlled in Table 6. I also use US occupational measures which are derived corresponding to each education level (i.e., high school degree, some college, and college degree or above). Table 10 presents the estimates from this exercise for the three generations.[32] Educational attainment appears to positively and significantly predict an individual's likelihood to work in professional jobs. The estimate on origin occupational wage stays positive and significant. The calculated marginal effects of a one standard deviation increase in origin wage display a declining trend over generations. Compared to column 3 of Table 6, the marginal effects are slightly smaller for all three generations, while the interpretations of the coefficients are not perfectly comparable.

To provide a direct comparison to the baseline results, I also run the conditional logit regressions on individuals with a high school degree, some college, and a bachelor's degree separately. The estimates suggest a negative effect of origin wage on non-college graduates, and the effect increases with educational attainment.[33] While some channels proposed in Section 3 (Conceptual Framework) predict a negative connection between origin wage and US occupation, the possibility that the negative coefficients reflect the constraints in the occupation choices available to lower-skilled workers cannot be ruled out. Relative to their higher educated counterparts, these people are more likely to work in jobs that require less training but are also associated with lower income. Such jobs may be at the lower end of the country-of-origin wage distribution as well. Notably, the regression results reveal a positive weight placed on hours of work among people without a college degree, regardless of generation, and a positive weight on the likelihood of unemployment among first-generation non-college graduates and native-born high school graduates. These findings controvert theoretical predictions (Boskin 1974; Dolton et al. 1989; Arcidiacono 2004) and may indicate a restricted choice set. Therefore, the nested logit model may be a more appropriate choice in this setting. To conclude, when workers of all education levels seek employment among the set of jobs that are potentially accessible to them, they appear more likely to take up jobs that were better-paid in their countries of origin.

9 Conclusion and discussion

This paper finds that occupational wage structures in origin societies are predictive of the job selection of immigrants and their descendants in the USA. *Ceteris paribus*, individuals are more likely to take up a job in the USA that was associated with higher monetary earnings in their country of origin. The weight assigned to the origin wage declines monotonically as individuals assimilate into the host society; i.e., spend more time in the USA or move to later immigrant generations. Even so, occupational patterns of later generations retain a distinctive mark from the labor market of their ethnic origins. The positive connection between the origin wage and the US

[32] Since the full sample is much larger now, I randomly select around 3000 individuals from each ethnic origin among the first generation and the native-born Americans for the origins with the number of observations that well exceeds 3000 given computational capability.

[33] Appendix Table 13 displays the results.

job choice is less likely to result from unobserved advantage. Instead, information asymmetry between immigrants' skills and local labor demands may explain this connection. Misinformation implies a welfare loss due to a worker-job mismatch. Thus, policies that facilitate information circulation may help immigrants overcome assimilation hurdles and avoid the potential economic and social costs of suboptimal investment in human capital of the later generations.

In considering these results, a few limitations need to be addressed. First, given data restrictions, this paper cannot differentiate all the possible mechanisms listed in Section 3 (Conceptual Framework). The test on the relationship between individuals' actual income and the origin wage of their chosen job, combined with a robustness check using the current wage rates in the country of origin, provide suggestive evidence that potential return migration may explain the empirical patterns. However, a formal test of this relationship can prove challenging. The current dataset does not provide any information on immigrants' intention to return home. Factors that may predict immigrants' propensity to return, such as years in the USA, refugee status, and economic and political factors in the origin country (Rooth and Saarela 2007; Dustmann et al. 2011), will inevitably capture other unobserved variations. Similarly, while social norms and cultural attitudes remain a plausible explanation, a clean proxy for cultural intensity to test this hypothesis is, again, difficult to obtain. If data on international scales of occupational status become available, one relevant question for future research is whether occupational prestige in the country of origin predicts US occupational choice, over and above the origin-country wage. Conditional on wage, prestige has the benefit of being relatively uncontaminated by all the possible labor-related channels. If origin occupational prestige is predictive of occupational choice that would be supportive of an interpretation that relies on cultural norms.

Second, this paper treats occupational wage distribution as exogenous. Migration may impact the wage structure of both the receiving country and sending country (Card 2001), and wages might increase in a country in occupations that observe a massive outflow of workers. As a result, those occupations may be precisely the ones overrepresented among immigrants in the USA. Nevertheless, dramatic wage changes can only happen in countries where the emigration rate (in an occupation) is large enough. Since occupational categories used in this paper are broadly defined, and since the emigration rates of countries in the sample are not high overall, the dynamic origin wage is less of a concern. Moreover, the origin wage is insignificant in most specifications when the origin fixed effects are controlled for; this implies a lack of temporal variation.

Lastly, ethnic enclaves and social networks serve as a critical medium for information flows. Social networks may provide immigrants with access to labor market niches and acquisition of new skills, possibly through information sharing. At the same time, ethnic enclaves may also limit employment opportunities if residents are isolated from information about the broader labor market (Warman 2007; Abramitzky and Boustan 2017). How pre-migration market experience interacts with ethnic networks in the USA is essential to the understanding of immigrants' search for employment, occupational choices, and assimilation. Though the current datasets do not allow distinguishing between the endogenous choices of occupation

and location, this relationship generates an important policy question worth future research attention.

Acknowledgments I thank Julie Cullen, Catalina Amuedo-Dorantes, Ernest Boffy-Ramirez, Oana Tocoian, Madeline Zavodny, the editor, Klaus F. Zimmermann, and four anonymous referees for helpful comments. Any mistakes are my own.

Data availability Additional results and copies of the computer programs used to generate the results presented in the paper are available from the author.

Compliance with ethical standards

Conflict of interest The author declares that she has no conflict of interest.

Appendix

A.1 Group mean levels of parental characteristics

As the American Community Surveys (ACS) does not collect information on the parents of respondents, I use a grouping estimation method (Card et al. 2000; Blau et al. 2013) to estimate the socioeconomic backgrounds of the 1.5 generation and the native-born Americans in the sample. The group means are derived from the population surveyed in the earlier censuses.[34] I use factors that may affect children's human capital accumulation, such as educational attainment, family income, and the number of children.

The data matching is as follows. First, groups of "parents" are identified for each individual. For 1.5-generation immigrants, their "parents" would be first-generation immigrants from the same country of origin who were aged 20–40 in an immigrant's year of birth and who migrated to the USA no more than five years after the birth year. For second- or higher-generation immigrants, their "parents" would be those, either foreign-born or native-born, of the same ethnic origin who were aged 20–40 and lived in the USA in an immigrant's year of birth. To avoid the potential problems of incomplete education and differential mortality, I examine only individuals aged 20–60 as of the survey year. Then I calculate the mean levels of parental and family characteristics adjusted for age and gender for each group of "parents."[35] The summary statistics of parental characteristics that match to my sample are reported in Table 1.

[34]Four censuses are used as follows: 1960 1% sample, 1970 1% metro sample, 1980 5% sample, and 1990 5% sample.

[35]The estimated group mean level is obtained by regressing the variable of interest on age, female, a year-of-survey dummy, and a full set of origin dummies. The estimated group mean level for a certain origin is the predicted value for a 40-year-old male from that nation surveyed in 1980. Besides, the number of children is estimated from only the group of individuals who had children born to them.

Admittedly, there is slippage in this grouping estimation method. The average parental socioeconomic characteristics of the highly educated sample are likely to be higher than the adjusted group mean levels obtained through the described approach due to positive human capital transmission (Solon 1992). Because of the substantial difference in educational attainment and income across racial and ethnic boundaries, the parental socioeconomic characteristics may be underestimated by different degrees for different ethnic groups. Nevertheless, the group mean levels capture the discrepancy in the human capital across ethnicities, which is an essential input in the formation of one's human capital (Borjas 1995b) and is an indicator of the ethnic human capital that helps individuals find employment (Fernandez and Fogli 2009).

A.2 Occupation categorization

The 383 three-digit coded occupations in the 1990 Census Bureau occupational classification scheme are categorized into 23 categories as follows: (1) general executives and managers such as legislators, chief executives, public administrators, and mail superintendent; (2) management-related occupations such as accountants, underwriters, and personnel specialists; (3) architects and engineers; (4) scientists and professors, including mathematicians, social scientists, and natural scientists; (5) doctors and lawyers; (6) health assessors, teachers, and librarians such as registered nurses, therapists, and secondary school teachers; (7) social workers such as recreation workers, clergy, and religious workers; (8) writers, artists, entertainers, and athletes; (9) engineering and science technicians such as electrical technicians, cartographers, and airplane pilots; (10) health and legal technicians such as practical nurses, dental hygienists, and paralegals; (11) sales representatives such as insurance agents, advertising agents, and sales engineers; (12) sales clerks such as cashiers, retail sales clerks, and street vendors; (13) office clerks and health service workers such as secretaries, interviewers, and dental assistants; (14) administrative support workers such as office supervisors, computer operators, and expediting clerks; (15) protective service workers such as firefighters, police, and sheriffs; (16) household and other service workers such as housekeepers, cooks, and janitors; (17) farming, forestry, and fishing occupations; (18) mechanics and system operators such as automobile mechanics, aircraft mechanics, and power plant operators; (19) repairers and precision workers such as office machine repairer and miners; (20) construction trades and craftsmen such as concrete and cement workers, engravers, and bakers; (21) heavy machinery operators such as ship crews, locomotive operators, and crane operators; (22) small machine operators and drivers such as printing machine operators, sawyers and bus drivers; and (23) laborers such as construction laborers and stevedores.

A.3 Other origin characteristics

The origin characteristics data are from various sources. The information for democracy status, state of war, and share of labor force in agriculture is acquired from

the Wejnert's Nations, Development, and Democracy Dataset from ICPSR. The per capita GDP data are from GDP and Per Capita GDP by Angus Maddison. The real GDP per capita is adjusted for purchasing power parity and expressed in 1990 International Geary-Khamis dollars. The Gini coefficient data are obtained from World Bank Open Data. The educational attainment data are from the Barro-Lee dataset. Distance to the USA is calculated as the number of air kilometers between the home country's largest city and the nearest US gateway (Los Angeles, Miami, or New York) using *www.timeanddate.com.* The information about nations' official languages is from *en.wikipedia.org/wiki/List_of_official_languages* .

A.4 Knowledge of the World of Work questions

The Knowledge of World of Work (KWW) survey, a part of the National Longitudinal Survey of Youth 1979 Children and Young Adults (NLSY79 Child/YA), is conducted on individuals aged 14 to 24. In the years 1994, 1996, and 1998, the NLSY79 Child/YA asked the young adults a series of questions concerning commonly held jobs. I focus on a set of eight questions regarding occupational earnings. Each question asks respondents to pick one out of two occupations that they think offers a higher wage. The respondents are also allowed to choose "Don't know" or "Refuse to answer." For example, one such question is "Who do you think earns more in a year? A person who is (A) a high school teacher or (B) a janitor."

Specifically, the questions compare: (1) automobile mechanic [mechanics/system operators][36] and electrician [construction trades]; (2) medical doctor [doctors/lawyers] and lawyer [doctors/lawyers]; (3) aeronautical engineer [architects/engineers] and medical doctor [doctors/lawyers]; (4) grocery store clerk [sales clerks] and truck driver [machine operators/drivers]; (5) unskilled laborer in mill [laborers] and unskilled laborer in factory [laborers]; (6) lawyer [doctors/lawyers] and high school teacher [health assessors/teachers]; (7) high school teacher [health assessors/teachers] and janitor [household/other service]; and (8) police officer [protective service] and janitor [household/other service]. Since the second and fifth questions contrast a pair of occupations that are in the same category, I exclude these two questions from the analysis.

Appendix Table 12 presents the summary statistics of the share of respondents of each ethnic origin selecting a certain occupation mentioned in the KWW questions.[37] Noticeably, in some pairs of occupations, such as those in the latter four questions, one occupation is predominantly selected by individuals of all ancestries. Yet variation exists across origins.

[36]The occupational category that a narrowly defined job falls in is reported in the brackets.

[37]I construct the adjusted share of individuals from an origin country choosing a certain option as follows: for each option in every question, I generate a binary indicator of the choice of a respondent and then regress the indicator on a set of origin dummy variables controlling for gender, age, and the year of survey. Accordingly, the coefficient on an origin dummy represents the propensity of 16-year-old males of the origin country selecting that option in 1994.

A.5 Tables

Table 11 List of origins and number of observations

Origin	First Gen.	1.5 Gen.	Native-born
Argentina	570	81	205
Australia	627	49	136
Austria	307	71	1476 (2956)
Belgium	418	53	1345
Bolivia	224	30	65
Brazil	999	65	111
Canada	611	298	577
Chile	213	20	31
China	1278 (21,398)	1401	1331
Colombia	438	89	90
Costa Rica	279	46	112
Denmark	282	92	1240 (4227)
Dominican Republic	928	90	329
Egypt	1067	123	171
El Salvador	1341	130	189
Finland	227	45	1480 (2483)
France*	364	69	289
Germany*	1452 (4354)	3801	1578 (179,277)
Greece*	194	43	251
Guatemala	661	61	99
Guyana	601	55	48
Haiti	599	32	32
Honduras	531	46	97
Hungary	439	74	1280 (4200)
India	1523 (37,021)	1286	1489 (1999)
Iran	438	24	70
Ireland*	154	260	262
Italy*	1431	1182	1444 (79,531)
Japan	1515 (3318)	771	1417 (5109)
Jordan	639	24	50
Lithuania	161	31	753
Mexico*	1564 (11,950)	1815	1469 (14,445)
Netherlands	564	168	730
Nicaragua	418	43	85
Nigeria*	954	32	62
Norway	345	213	1537 (17,767)

Table 11 (continued)

Origin	First Gen.	1.5 Gen.	Native-born
Pakistan	638	75	105
Peru	1492 (2051)	108	318
Philippines*	1481 (12,855)	1553	1452 (3616)
Poland*	558	139	432
Portugal*	897	217	1479 (3059)
Romania	1502	71	959
Russia*	1246 (3681)	175	1250 (10,241)
Sweden	368	131	1115
Switzerland	82	23	270
Thailand	533	87	139
Trinidad and Tobago	441	28	47
Turkey	627	53	173
UK*	1468 (9385)	3363	1530 (154,128)

Reported are the numbers of observations from each origin in the sample. The numbers in the parentheses are the original numbers of observations before trimming. "*" denotes origins included in the Knowledge of the World of Work dataset

Table 12 Subgroups of first-generation immigrants

	Migrated after 18		Naturalized citizens	
Variable	(1)	(2)	(3)	(4)
Origin occupational wage	.115***	.128***	.096***	.121 ***
	(.017)	(.015)	(.020)	(.017)
US occupational wage	.948***	.987***	.849***	.884***
	(.071)	(.063)	(.065)	(.065)
Marginal effect	.482	.545	.400	.504
	[11.6%]	[13.0%]	[9.61%]	[12.2%]
No. of observations	662,003	662,003	414,570	414,570
Log-likelihood/10^4	−9.08	−8.92	−5.56	−5.51
Individual characteristics	Yes	Yes	Yes	Yes
Ethnic networks	Yes	Yes	Yes	Yes
Origin characteristics	No	Yes	No	Yes
Origin fixed effects	No	No	No	No

*Significant at 10%; **Significant at 5%; ***Significant at 1%. Regressions are estimated by the conditional logit model, weighted by the ACS person weight. The first-generation immigrants are examined. The dependent variable is an indicator of choice among the 23 occupation categories. Robust standard errors in parentheses are clustered at origin level. Non-alternative-specific variables are introduced to the regressions by interacting them with the six occupational attributes

Table 13 Origin wage distribution and occupational choices by education: conditional logit

	First generation			1.5 generation			Native-born		
	High school	Some college	Bachelor degree	High school	Some college	Bachelor degree	High school	Some college	Bachelor degree
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Origin occupational wage	−.347***	−.032	.135***	−.393***	−.138***	.098***	−.285***	−.077**	.049***
	(.054)	(.024)	(.014)	(.049)	(.028)	(.017)	(.045)	(.030)	(.018)
US occupational wage	.236***	.250 ***	1.023***	.348***	.522***	.919***	.192***	.430***	.863***
	(.064)	(.048)	(.071)	(.044)	(.051)	(.050)	(.058)	(.045)	(.052)
Years of experience	−.084*	−.049	.195***	−.102***	−.283 ***	.192***	−.020	−.186 ***	.162***
	(.048)	(.056)	(.038)	(.031)	(.068)	(.039)	(.042)	(.057)	(.036)
Hours of working	.178***	.179 ***	−.731***	.164***	.172***	−.514***	.223 ***	.170***	−.388***
	(.023)	(.028)	(.066)	(.021)	(.032)	(.063)	(.031)	(.035)	(.045)
Education required	−.417***	−.374 ***	−.223***	−.284***	−.268***	−.272***	−.295 ***	−.324***	−.230***
	(.055)	(.053)	(.045)	(.065)	(.047)	(.055)	(.049)	(.051)	(.044)
% unemployed	.191***	.087**	−.363***	.023	−.135***	−.590***	.038	−.089*	−.567***
	(.044)	(.040)	(.033)	(.050)	(.047)	(.061)	(.059)	(.046)	(.046)
Marginal effect	−1.20	−.129	.559	−1.35	.532	.402	−1.04	−.305	.197
	[28.4%]	[3.03%]	[13.7%]	[31.5%]	[12.4%]	[9.82%]	[24.0%]	[7.09%]	[4.79%]
No. of observations	268,488	299,040	785,784	200,579	273,184	429,915	203,751	343,695	642,869

Table 13 (continued)

	First generation			1.5 generation			Native-born		
	High school	Some college	Bachelor degree	High school	Some college	Bachelor degree	High school	Some college	Bachelor degree
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
No. of individuals	11,836	13,088	34,286	8788	11,934	18,736	8956	15,043	28,007
Log-likelihood/10^4	−4.20	−4.47	−10.1	−2.99	−3.95	−5.10	−2.87	−4.65	−7.56

*Significant at 10%; **Significant at 5%; ***Significant at 1%. Regressions are estimated using the conditional logit model on separate samples by education and generation, weighted by the ACS person weight. The dependent variable is an indicator of choice among the 23 occupation categories. Robust standard errors in parentheses are clustered at origin level. All US job characteristics are derived by education level. Individual, parental, origin characteristics, and ethnic networks are controlled for in all the specifications. Non-alternative-specific variables are introduced to the regressions by interacting them with the six occupational attributes. For first-generation immigrants, parental income, education, and the number of siblings are not included in the control set. The reported marginal effects are the average percentage point change in choosing one occupational category given a one SD increase in the standardized occupational wage of that category in the origin country. The marginal effects are in percentage points, with percentage changes reported in brackets

References

Abramitzky R, Boustan LP, Eriksson K (2014) A nation of immigrants: assimilation and economic outcomes in the age of mass migration. J Polit Econ 122(3):467–506

Abramitzky R, Boustan L (2017) Immigration in american economic history. J Econ Lit 55(4):1311–45

Abramitzky R, Boustan LP, Eriksson K (2020) Do immigrants assimilate more slowly today than in the past? American Economic Review-Insights 2(1):125–41

Arcidiacono P (2004) Ability sorting and the returns to college major. J Econ 121(1):343–375

Betts JR, Lofstrom M (2000) The educational attainment of immigrants: trends and implications. In: Borjas GJ (ed) Issues in the economics of immigration. University of Chicago Press, pp 51–116

Bisin A, Verdier T (2000) Beyond the melting pot: cultural transmission, marriage, and the evolution of ethnic and religious traits. Quart J Econ 115(3):955–988

Blau FD, Kahn LM, Papps KL (2011) Gender, source country characteristics, and labor market assimilation among immigrants. Rev Econ Stat 93(1):43–58

Blau F, Kahn L, Liu A, Papps K (2013) The transmission of women's fertility, human capital, and work orientation across immigrant generations. J Popul Econ 26(2):405–435

Borjas G (1987) Self-selection and the earnings of immigrants. Am Econ Rev 77(4):531–553

Borjas GJ (1992) Ethnic capital and intergenerational mobility. Quart J Econ 107(1):123–150

Borjas GJ (1995a) Assimilation and changes in cohort quality revisited: what happened to immigrant earnings in the 1980s? J Labor Econ 13(2):201–245

Borjas GJ (1995b) Ethnicity, neighborhoods, and human-capital externalities. Amer Econ Rev 85(3):365–390

Boskin MJ (1974) A conditional logit model of occupational choice. J Polit Econ 82(2):389–398

Card D, DiNardo J, Estes E (2000) The more things change: immigrants and the children of immigrants in the 1940s, the 1970s, and the 1990s. In: Borjas GJ (ed) Issues in the economics of immigration. University of Chicago Press, pp 227–270

Card D (2001) Immigrant inflows, native outflows, and the local labor market impacts of higher immigration. J Labor Econ 19(1):22–64

Chiswick BR (1978) The effect of americanization on the earnings of foreign-born men. J Polit Econ 86(5):897–921

Chiswick BR (1983) An analysis of the earnings and employment of asian-american men. J Labor Econ 1(2):197–214

Chiswick BR, Taengnoi S (2007) Occupational choice of high skilled immigrants in the united states. Int Migr 45(5):3–34

Constant AF, Zimmermann KF (2003) Occupational choice across generations. Appl Econ Q 49(4):299–317

Dolton PJ, Makepeace GH, van der Klaauw WH (1989) Occupational choice and earnings determination: the role of sample selection and non-pecuniary factors. Oxf Econ Pap 41(3):573–594

Dustmann C, Fadlon I, Weiss Y (2011) Return migration, human capital accumulation and the brain drain. J Dev Econ 95(1):58–67

Dustmann C, Preston I (2012) Comment: Estimating the effect of immigration on wages. J Eur Econ Assoc 10(1):216–223

Dustmann C, Schonberg U, Stuhler J (2016) The impact of immigration: why do studies reach such different results? J Econ Perspect 30(4):31–56

Edin P-A, Fredriksson P, Aslund O (2003) Ethnic enclaves and the economic success of immigrants - evidence from a natural experiment. Quart J Econ 118(1):329–357

Eriksson K (2019) Ethnic enclaves and immigrant outcomes: Norwegian immigrants during the Age of Mass Migration. Eur Rev Econ Hist 24(3):427–446

Fernandez R, Fogli A (2009) Culture: an empirical investigation of beliefs, work, and fertility. Amer Econ J Macroecon 1(1):146–77

Giuliano P (2007) Living arrangements in western europe: does cultural origin matter? J Eur Econ Assoc 5(5):927–952

Hanson GH, Liu C (2017) High-skilled immigration and the comparative advantage of foreign-born workers across us occupations. In: Hanson GH, Kerr WR, Turner S (eds) High-skilled migration to the United States and its economic consequences. University of Chicago Press, pp 7–40

Keith K, McWilliams A (1999) The returns to mobility and job search by gender. Ind Labor Relat Rev 52(3):460–477

Lafortune J, Tessada J (2012) Smooth(er) landing? The dynamic role of networks in the location and occupational choice of immigrants, Working Paper, Pontificia Universidad Catolica de Chile
Lewer JJ, den Berg HV (2008) A gravity model of immigration. Econ Lett 99(1):164–167
Munshi K (2003) Networks in the modern economy: Mexican migrants in the u. s. labor market. Quart J Econ 118(2):549–599
Peri G, Sparber C (2011) Highly educated immigrants and native occupational choice. Ind Relat J Econ Soc 50(3):385–411
Propper C, Reenen JV (2010) Can pay regulation kill? Panel data evidence on the effect of labor markets on hospital performance. J Polit Econ 118(2):222–273
Rooth D-O, Saarela J (2007) Selection in migration and return migration: evidence from micro data. Econ Lett 94(1):90–95
Ryoo J, Rosen S (2004) The engineering labor market. J Polit Econ 112(S1):S110–S140
Schaeffer PV (1995) The work effort and the consumption of immigrants as a function of their assimilation. Int Econ Rev 36(3):625–642
Solon G (1992) Intergenerational income mobility in the United States. Amer Econ Rev 82(3):393–408
Spitzer Y, Zimran A (2018) Migrant self-selection: anthropometric evidence from the mass migration of italians to the united states, 1907-1925. J Dev Econ 134:226–247
Warman C (2007) Ethnic enclaves and immigrant earnings growth. Can J Econ 40(2):401–422
Wobmann L (2003) Schooling resources, educational institutions and student performance: the international evidence. Oxf Bull Econ Stat 65(2):117–170
Zhan C (2015) Money v.s. prestige: cultural attitudes and occupational choices. Labour Econ 32:44–56