UNIVERSITY OF BONN

RESEARCH MODULE: APPLIED MICROECONOMICS

Native - Immigrant Gaps in Occupational Choices: the Case of Germany

Anh-Thu Duong, Nicolas Fajardo, Joran Poortstra, Shu-Ya Wei

February 5, 2021

CONTENTS

ABSTRACT

Regarding employment opportunities, ethnicity differences are among the most discussed questions of labor economists. Using data from the 1984-2018 German Socio-Economic Panel, we apply non-structural analysis to examine the occupational status gaps between native Germans and immigrants. In this paper, we develop two different occupational classification methods based on the skill-cell approach: the first classification relies on the ISCO-08 job skills level, whereas the second approach bases on the level of language proficiency as a job requirement. However, due to the inherent limitations of our modeling method, which fails to capture the heterogeneity of the assimilation process, we could not find large differences existing between native-born Germans and both the first- and the second-generation of immigrants after controlling for one's education, experience, and other socio-demographic characteristics.

1. INTRODUCTION

The labor market impact of past, current, and future migration influxes has been raised as an integral part of the discussion on both the scientific and policy-making stances. For example, Bodvarsson et al. (2014) state that the economics profession today understands migration as "primarily driven by a desire to maximize one's return to human capital investment, [implying that] people respond to spatial differences in labor market opportunities by migrating if those opportunities dominate the costs of relocation". In this paper, our main objective is to study the occupational decisions of both immigrants and natives from a supply-side perspective using the German Socio-Economic Panel (SOEP). More specifically, with help of the O*NET database in combination with the SOEP data, we analyze to what extent measurable socio-demographic characteristics can explain the labor market choices of workers in Germany, with respect to skill level and language intensity of occupation, when workers are assumed as rational decision-makers driven by purely utilitarian considerations.

Whereas standard economic theory suggests that an increase in the supply of labor would decrease wages and possibly force native-born workers to compete directly with immigrant workers, empirical evidence shows a completely different picture. For instance, Chiswick (1978) finds that for immigrants, language ability and time spent in the destination country are the most important determinants of earnings, questioning the perfect substitutability commonly assumed to exist between both groups. Peri and Sparber (2009) focus their study on less-educated immigrants and comparably educated native counterparts, yet find a clear trend regarding the task specialization between immigrant and native-born workers, who are imperfect substitutes in the labor market of the United States. Amuedo-Dorantes and De La Rica (2011) also explore this topic in the case of Spain and observe that native workers switch to more language-complex jobs as a response to immigrants entering the labor market. In general, empirical literature remarks that while immigrant workers choose jobs that are intensive in physical labor or need *low language skills*, native-born workers specialize in jobs that require relatively more language abilities; defined as having a marked intensity of communication and interaction tasks (*high language skill jobs*).

In addition, Cunha and Heckman (2007) show that families play a crucial role in shaping individual skills and, thus, create the observed ability gaps between different socioeconomic groups, inasmuch as they seem to appear at very early ages. Therefore, relying on those considerations, we conjecture that the individual migration background has a strong impact on the occupational choice and its effects persist through several generations. Even for second-generation immigrants—children of direct immigrants who were born in Germany or immigrated at a very early age, and learned German (at least in school) as their first language—, the cultural assimilation into the German society and its economic consequences could still depend largely on their parents' social and cultural background and its inheritance patterns. Nonetheless, if certain characteristics like race and sex may indeed affect one's occupation because of discriminatory practices exerted by employers, it is also equally plausible that race and sex may affect one's preferences for different occupations, therefore, —taking into direct consideration our reduced-form approach— we are not able to unravel the differential impact of social discrimination and individual utility considerations on occupational choice, consequently, our results should be taken with caution.

Overall, our model does not allow a straightforward conclusion about the difference in employment opportunities between native German and immigrants. One possible reason is that our classification method is not able to get a clear country-specific picture of immigrants. Although we recognize that those of Turkish origin are one of the most disadvantaged groups of immigrants in Germany, evidence on occupational success gaps is not statistically significant. Another reason could be that the simple human capital accumulation model fails to capture the individual heterogeneity in a migration context. However, our findings suggest that the observed occupational choice is perfectly in line with basic human capital accumulation theoretical predictions: Together with an efficacious assimilation of newcomers into the host country, higher language proficiency, higher education, and relevant working experience are necessary for boosting upward occupational progress.

The paper continues as follows, first we describe the existing research on human capital accumulation of immigrants in the labor market and give a brief overview of the immigration history of Germany, highlighting important characteristics captured in our sample. Then, after a thorough description of our data, their specific variables, and both theoretical and empirical explanations of their adequacy, we detail our proposed estimation methodology and explain its qualities as a reduced-form analysis. Finally, we report our results and conclude with a short discussion.

2. LITERATURE REVIEW

In this section, we present the most relevant and discussed stylized facts on immigration. In the first part, we present the differences between natives and immigrants framed in the human capital theory. In the second part, we lay out a brief history of immigration in Germany and its perceived economic consequences.

2.1. Immigration and the Human Capital Accumulation

Human Capital Research has a long history focusing primarily on the returns of education. The central tenet of the theory is that "individuals and society derive economic benefits from investments in people. The investment feature of this suggestion significantly differentiates human capital expenditures from consumptive expenditures—those providing few benefits beyond immediate gratification" (Sweetland, 1996). In that regard, Fleischhauer (2007) establish three categories to classify the human capital literature: (i) measurement of the returns, (ii) production function, and (iii) human capital formation. Since we are focusing on the differences of occupational choice, rather than the actual effect of education between natives and immigrants, we will review how the human capital formation approach fits into the present study—acknowledging the existence of a vast literature on human capital accumulation as a driver of employment outcomes.

As a framework for theoretical developments on the determinants of occupation, we begin by reviewing basic human capital model. The most widely used specification of empirical earnings equations and the beginning point of our analysis is the Mincer equation:

$$ln(earnings) = \alpha + \rho_s schooling + \beta_0 experience + \beta_1 experience^2 + \varepsilon$$

This regression model is motivated by two conceptually different frameworks described in Mincer (1958, 1974). More specifically, Mincer (1958) model uses the principle of compensating differences to explain why individuals with different levels of schooling receive different earnings (have different occupation status) over their lifetimes. The model used by Mincer (1974) shows a linearly declining rate of post-school investment, assuming that the effects of schooling and experience are identical across people.

Regarding the assimilation process of immigrants, Miyar-Busto et al. (2020) highlight that the transferability of human capital using language skills is an important factor to help guest workers to access employment opportunities in Spain. In Fellini et al. (2018), it is shown that the return to post-secondary education of non-Western immigrants is significantly lower for their first job in the Italian labor market. On the contrary, the returns of other West-European immigrants depend on the country of origin, suggesting that the gap in the transferability and quality of skills are scarcely relevant in a strongly segmented labor market. Using longitudinal Belgian data, Baert et al. (2016) shows there are native-immigrant gaps in educational attainments and school-to-work transitions. To be more specific, immigrant students are less likely to transition into work successfully, and female immigrants' performance gaps are substantially larger.

In the case of Germany, Algan et al. (2010) find substantial differences between first-generation immigrants and their native counterparts, as well as within the origin countries of immigrants. Hartmann (2014) suggests that there is a big gap between native-born Germans and second-generation Turkish. However, while the results for second-generation Turkish males show that their differences can be explained entirely by schooling, the root causes of the disadvantage between second-generation Turkish women and native German women are more complex. Diehl and Granato (2018) show that immigrants have jobs with lower occupational status than native Germans, even if they have equivalent educational degrees. The results of Basilio et al. (2017) reveal that, in general, education and working experience accumulated in home countries receive significantly lower returns than human capital obtained in Germany. They also find evidence for heterogeneity in the returns to human capital of immigrants across countries of origin.

2.2. An overview of immigration influx to Germany

Despite Germany being traditionally considered as a no migration country, after World War II the country has received important inflows of migrants. According to Bauer et al. (2005) there were four phases of migration streams to Germany: (i) war adjustment [1945–54] (ii) manpower recruitment [1955–73] (iii) consolidation or restrained migration [1974–88] (iv) and the dissolution of socialism and its aftermath [1988-98]. Following Hess and Green (2016) and Brücker et al. (2020), it is also necessary to add two recent immigrant groups: (v) post 1998 Federal Election wave [1998-2014] (vi) refugee wave [since 2015]. In general, all waves have been characterized by a systematic distinction of international immigration groups based on their citizenship status, that is, on one hand, people of German descent —ethnic German immigrants from Central and Eastern Europe—, and foreign workers, refugees and asylum seekers in the other, whose right to immigrate to Germany has been subject to strict regulation. According to the Statistisches Bundesamt (2020), Germany's Federal Statistical Office, around 13 percent of people living in Germany have a migration background and 12 percent are of foreign descent. Of these 21 million, the three largest ethnic groups are Turkish (13 percent), Polish (10 percent), and Russian (7 percent) with either a direct or indirect migration background.

In this paper we will not only focus on first-generation (*direct*) immigrants but also on second-generation (*indirect*) immigrants. The differences persist between second-generation immigrants and natives despite that second-generation immigrants have, in theory, the same access to and possibility of education and development as their native counterparts. Furthermore, language proficiency is also to be expected higher for second-generation immigrants.

For first-generation immigrants, the difference in occupational choice with respect to the *language skill of jobs* plays an important role in employment outcomes. Using the German Labor Force Survey (DE-LFS) on low-skilled natives and immigrants, Sebastian and Ulceluse (2019) conclude that in the context of increasing labor supply of immigrants, natives would shift their task supply and provide more communication relative to manual tasks, which is in line with Peri and Sparber (2009) findings. Moreover, the timing of migration is an endogenous choice variable. It is influenced by unobserved ability and other unmeasured wage factors, leading to an "ability bias". The problem is further complicated by individual heterogeneity in the slopes of earnings progression, which may be jointly determined with the timing of migration. Anticipated post-migration wage growth may influence the decision of when to immigrate, potentially causing "slope-heterogeneity bias" (Jain & Sabirianova Peter, 2017). Algan et al. (2010) run a simple regression with earnings as the dependent variable and controlling for basic characteristics and find that first-generation immigrant men earn less than their native counterparts, with huge differences between the groups of countries of origin. In short, first-generation immigrants choose whether to immigrate (to Germany) or not and thus a selection bias may arise, a problem that is not strongly relevant with second-generation immigrants.[1] Since we use a panel data set —meaning repeated observations over time for multiple individuals— macroeconomic, political, and societal factors may also play an important role in the decision to immigrate or not. One example is the Syrian refugee crisis of the mid-2010s. In our sample, a large increase in immigrants from Syria is visible in the years thereafter. These refugees, compared to (economic) immigrants, have not immigrated voluntarily to Germany but had to flee their home country.

However, second-generation immigrants are also not homogeneous. In particular, cultural differences of parents of second-generation immigrants or the location they grow up in can affect their final occupational choices. Hartmann (2014, 2016) puts the focus on the differences between native-born Germans and second-generation Turks. There are two main findings; first, the degree of assimilation of second-generation Turkish is low and they experience higher unemployment, which holds for their entire early employment career and is not just a temporary phenomenon at some stage of their career. Their lower educational qualifications and lower rates of vocational training are the main reasons for these struggles towards middle-class careers. Second, there are greater difficulties for second-generation Turkish women to assimilate into the middle class than for second-generation Turkish men. Furthermore, an individual's ethnic background can affect job seekers in the labor market. Sürig and Wilmes (2015) examine institutional discrimination by enterprises. It is shown that structural elements such as job skills, work content and responsibilities are not the only determinant of success in the workplace. Ethnically, culturally, or socially motivated discrimination also affects one's job experiences. Also, there are remarkable differences among genders; while nearly two-third of Turkish-background males reported that they were affected by hostilities while looking for a job, half of the women of Turkish origin had such an experience.

3. DATA AND VARIABLES

We conduct our analysis based on consolidated data from the German Socio-Economic Panel (SOEP). The major advantage of the SOEP data set is that it allows us to link the inter-temporal adults' career decisions from 1984 to 2018 with their generated biographical information, displaying not only some time-invariant individual characteristics —such as migration background or country of origin— but also precise information about their parents. Moreover, given the considerable data consistency, we are also able to measure variables in a very detailed fashion like educational attainment, work experience, marital status, etc.[2] That said, this section continues the discussion about the generation and cleaning processes applied to raw SOEP data as well as the main patterns found in some selected variables, which in turn were chosen by their theoretical or empirical relevance within the migration and human capital literature. More detailed plots and distribution regularities can be found in section A in the appendix.

[1] Although we could argue that observed second-generation immigrants also decided to stay, which means that those more apt have gone to their family's origin or other more prosperous countries, which is unlikely the case of Germany.

[2] For example, we find in other studies that, as work experience is often not accurately measured, researchers proxy it as age minus years of education minus 5. In this case, we don't have to explicitly recur to such an *ad hoc* imputation method.

MAJOR GROUPS	CLASSIF.	SKILL LEVEL
1. Senior officials and managers	*Professional*	3 + 4
2. Professionals		4
3. Technicians, associate professionals	*Technical*	3
4. Clerks	*White-collar*	2
5. Service, sales workers		2
6. Skilled agricultural and fishery workers	*Blue-collar*	2
7. Craft and related workers		2
8. Plant and machine operators and assemblers		2
9. Elementary occupations	*Elementary*	1
0. Armed forces		1 + 2 + 4

Table (1) Occupational categories of ISCO-08

3.1. Migration background

According to the SOEP, there are three different migration backgrounds: (i) native Germans with no migrant background, (ii) direct (*first-generation*) immigrants and (iii) indirect (*second-generation*) immigrants. More precisely, the children of first-generation immigrants, denominated second-generation immigrants are defined as those individuals who were born in Germany and whose mother and father were born abroad. However, one of the drawbacks of using the migration background variable reported by the SOEP is that in cases where one of the parents is native German and the other is a foreigner, it cannot always correctly identify an individual as having a relevant migration background. For example, if a German marries a Turkish, their children could be considered native Germans despite one of their parents being an immigrant. Thus, the migration background [*migback*] variable slightly underestimates the number of indirect immigrants. In those cases, we acknowledge that issue by changing their status to "indirect immigrant", if and only if the origin information of both parents is available and at least one of the parents is a foreigner.

Moreover, by construction, the SOEP is not able to identify the migrant status of an individual before 1950, hence, only persons being born after 1949 are kept in the sample, as well as direct immigrants with an identifiable year of arrival to Germany, all other deviations were dropped out. Finally, under the umbrella of the human capital approach, we are able to consider 1.5 generation immigrants as full second-generation (indirect) immigrants, since these are foreign-born individuals with foreign-born parents but arrived in Germany before 7 years old, implying that they face the German "acculturation shock" at a very early age, attend to the same pre-school educational system and learn German at a faster rate than comparable older direct immigrants (Schüller, 2012). Eventually, even when we changed their migration status, we still retain the length of their stay in Germany and control for it in later stages.

3.2. Occupation classification of ISCO-08

To account for occupational decisions of both native and immigrants, we classify all the International Standard Classification of Occupations (ISCO) reported jobs into five groups, from low to high skill: (i) Elementary, (ii) Blue-collar, (iii) White-collar, (iv) Technical and (v) Professional. This distinction is made taking into account the skill-based structure defined by the ISCO in its 2008 version, where skill is "defined as the ability to carry out the tasks and duties of a given job (...) [Measured in] two dimensions (...) *skill level* and *skill specialization*." For instance, the skill level is the primary reference for the classification of *Major Groups* and measures the nature of the work performed, the level of formal education and the amount of informal on-the job-training (International Labour Office, 2012). Therefore, such a classification fits naturally into the conceptualization of the Human Capital Model previously characterized. Although white-collar and blue-collar jobs are classified into the same skill level, we intentionally differentiate between the two choices to emphasize its differential sectoral nature —service and manufacture respectively— and account explicitly for the fact that sectoral tastes and physical ability could also influence the individual occupational decisions (see Table 1). More specifically, to construct a harmonized ISCO code across the different waves, we use the Personal Generated Data [*pgen*] and its reported ISCO from both 1988 and 2008 versions, matching both variables using the crosswalk published by the International Labour Organization, and adding some imputed relations in order to cover all existing values of ISCO-88.[3] Since the SOEP sometimes reports both ISCO versions, the occupation variable is reclassified only in such a

[3] The decision rule was set to select the highest available unit, minor or sub-major group in ISCO-08 covering the raw descriptions provided by the ISCO-88 classification, therefore, a slight upward skill bias exists for all our calculations in 18 of 403 ISCO reclassified units. Note that units refer to the classification scheme and not the total individual observations that were modified. About the crosswalk, Kvetan (2014) comment that important changes between Major Groups occurred during the update, therefore it can be expected that ISCO Major Groups show —in the number of employed— [1, 3, 4, 8 and 9] a net decline [2 and 5] a net increase [6 and 7] relatively same numbers.

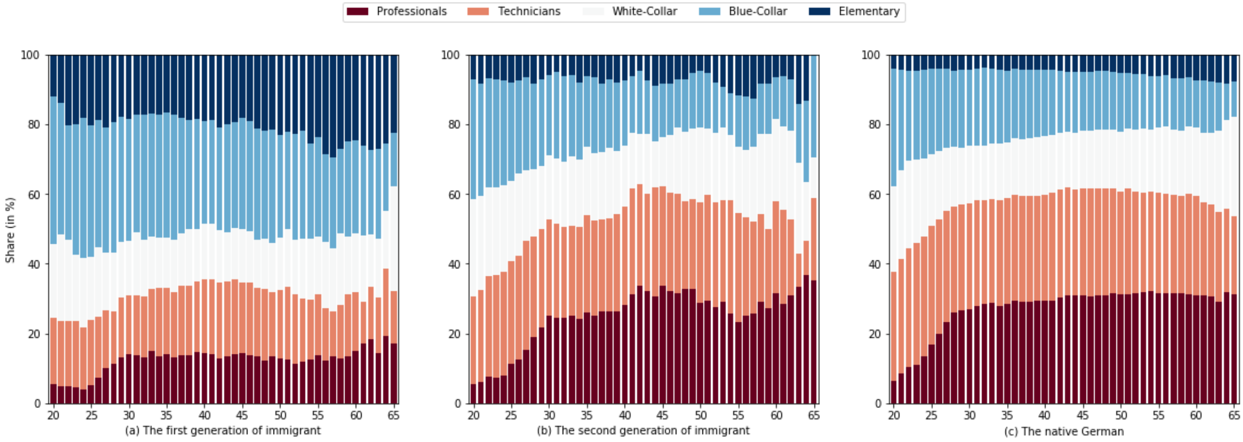


Figure (1) ISCO Skill occupational decisions by age and migration background

case where the ISCO-88 is available but the ISCO-08 is not. Finally, since not-employed observations do not report their associated 4-digit ISCO code, we are unable to classify them and, thus, drop them from the dataset.

In addition, following Dustmann (2004), it should be noted that career choices of the individuals educated in Germany are highly correlated with their educational attainment due to the intrinsic design of the German school system, because it imposes critical restrictions on the career choice of an individual at a very early age; placing pupils into different secondary schooling tracks at around the age of 10 years and almost determining their career track jointly with the school choice: The lower level *Hauptschule*, designed to prepare pupils for manual jobs; the intermediate *Realschule*, which prepares students for administrative and lower white-collar jobs and finally the upper-level *Gymnasium*, which prepares students for tertiary education.

Figure 1 depicts the age-occupation profiles of first (direct) and second (indirect) generation immigrants, and native-born Germans. In general, there are no evident differences between the distribution of occupations of both indirect immigrants and native Germans, in the two cases, the proportion of technicians and professionals increase and remain stable after the age of 30. However, when we look closely at the age-profiles of second-generation immigrants differentiated by ethnic groups, we notice an important gap between Germanic and Turkish individuals (see Figure 12), considering that while the percentage of Turkish being in professional and technical jobs remains below 50%, the proportion of indirect German immigrants joining these two occupations rise significantly from 30% at the age of 20 to around 70% between 30 and 50 years old. The latter suggests there is a considerable amount of heterogeneity between ethnic groups within second-generation immigrants.

3.3. Occupation classification based on Language Skill

The assimilation effect outlined by Chiswick (1978, 1979), in which migrants often lack skills specific to their destination country and thus face transaction costs[4] to "fully-transfer" their abilities into the new labor market, contemplates that, as time passes, migrants are able to build destination-specific skills at a decreasing rate. Such a "convexity" has two interesting properties: (i) allows immigrants to undertake finite human capital investments in their destination country primarily determined by their respective individual gains, (ii) the payoff time for such investments would be the greatest the earlier they are done, therefore, constituting age of arrival and years in the destination as reasonable predictors of direct migrants' occupational success. Consequently, besides using the job skill level defined by ISCO-08, we also categorize occupations based on the language skills needed to perform each respective job. This approach enables us to evaluate the assimilation prediction because it is assumed that, in absence of educational or linguistic transaction costs, there should be no difference in skill attainment, therefore setting the second-generation immigrants as an important "control" group for analyzing the accuracy of the model and investigate other kinds of factors determining the native-immigrant occupational gaps such as cultural tastes and family background persistence.

For that matter, we map the required linguistic skill of each ISCO-08 reported occupation, using the data provided by O*NET *à la* Chiswick and Miller (2010) in two different ways: (i) characterizing only the level and importance of linguistic knowledge required for each job, namely the "knowledge of the structure and content of the [corresponding] language including the meaning and spelling of words, rules of composition, and grammar." and (ii) aggregating the previously explained linguistic knowledge level and its importance for a specific job with the reported Writing, Speaking, Reading and Listening level and importance scales attached to it. In the end, both measures are indeed very similar, but we opt to choose the latter as it entails a much more realistic picture of the linguistic skills necessary to effectively perform a job. Additionally, two further aspects have to be taken into account. First, that level and importance measures are

4 Where these costs are primarily driven by language, cultural and institutional differences.

mapped through a survey conducted by the O*NET service in the United States and therefore reflect the idiosyncratic characteristics of the US labor market, which could be arguably very different from its German counterpart.[5] Second, that the construction process of the language skill variable heavily relies on the underlying ISCO-08 structure to maintain a certain degree of comparability, therefore a pure linguistic interpretation must be avoided.[6]

In Figure 2, we present the distribution of the second approach of linguistic skill classification (the aggregate of different aspects) in five quantiles by migration background. Occupational choices of native Germans are, as expected, stable over age. Note the analogous structures between Figure 1 and 2, meaning that our classification based on required language skills for jobs are highly correlated to the classification ISCO skills for jobs. We do perform some comparisons between ISCO-08 job skills and language skills in Table D and Figure 25. The consistent positive correlation across classification suggests that lower language proficiency could possibly downward one's occupational status. Moreover, observing the surprising similarity in the correlation of the language-skill requirement and job-skills requirement for the native Germans group, direct and indirect immigrants group (73%, 74% and 74% language skill level 5 for professionals, respectively), we can say that required linguistic skill, which native Germans are assumed to be totally fluent and have advantages, is not simply a daily communication skill but professional language skill as a job requirement.

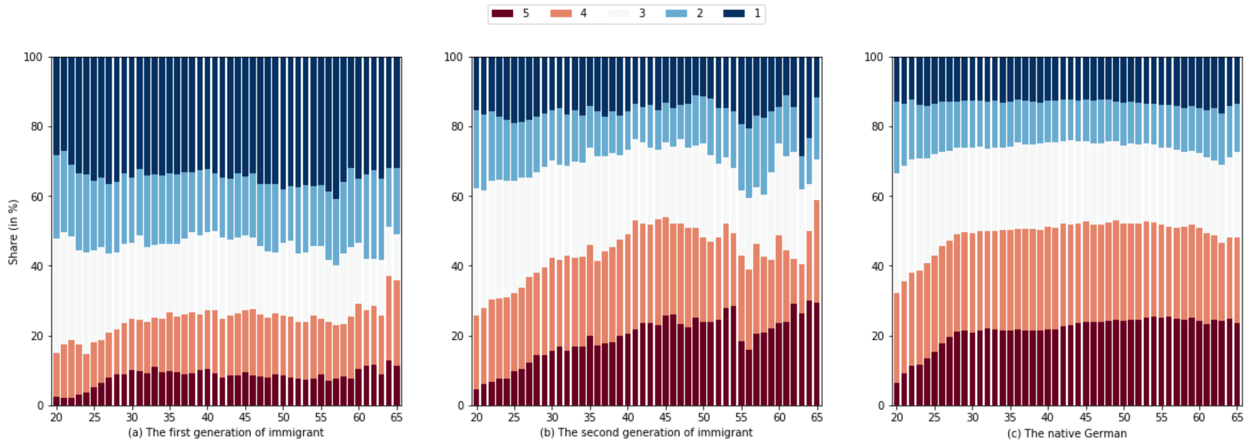


Figure (2) Language Skill occupational decisions by age and migration background

3.4. Cultural groups and Language Distance

Note that with the heterogeneity between immigrants from different origin countries but infeasible controlling at country-level differences given small sample restrictions, we explore the cultural diversity at group-level approach. A few studies have discussed that cultural values could shape individual preferences and impact the utility of choosing between different occupations, especially when comparing native and migrant populations. Examples include Beugelsdijk et al. (2017), who perform an extensive review of the rule of culture from an International Business perspective,[7] and argues that there is an important intra-country heterogeneity and therefore culture and country effects must be differentiated[8]. Nonetheless, as particular efforts to objectivize and measure distances on work-related values between countries remain somewhat arbitrary,[9] Ronen and Shenkar (2013) conduct a meta-study using different classifications within the International Business literature and group countries into cultural zones robust to measure uncertainty, resulting in 11 cultural clusters with different levels of cohesiveness (see figure 8). Yet, one drawback remains, only 96 from almost 150 countries reported

5 In our defense, both countries belong to the upper tail of world's GDPs and share comparable institutional arrangements

6 The exact procedure to construct the language-skill-based occupational classification goes as following: (a) since O*NET databases make use of the classification of the US Census Bureau, we use the crosswalk from O*NET-SOC-10 to ISCO-08 published by the Institute for Structural Research. (b) then, to handle the remaining missing data, we perform a stepped data imputation process taking advantage of the ISCO classification construction: (i) we reduce all 4-digit codes to a 3-digit form, and impute the missing observations to the median of the available values of the same 3-digit group, (ii) next, the 3-digit codes are further reduced to a 2-digit form and the same imputation operation is performed (iii) finally, since all ten (10) ISCO major groups are also classified by four (4) different skill-levels, we attributed the median of each skill group to the remaining missing observations. In the end, for the composite *language-interactivity* measure we aggregate the five-level indexes (linguistic knowledge, writing, speaking, reading, and listening) using a weighted by importance average, re-scale the resulting variable between 0 and 100, and cut the obtained results in five quantiles to resemble ISCO's five occupational sectors. For further details see section D listing the specific job units being moved across occupational groups between ISCO and Language Skill measures and figure 25 to compare changes between the raw classification density and the actual sample density conditional on migration background.

7 (i) view culture as a set of values that are shared in a given social group and distinguish this group from others and (ii) explicitly focus on differences between work-related values

8 cultural differences may be more region- than country-specific, (...) [meaning] that cultural values exhibit marked discrete jumps at the boundaries of these supra-national cultural zones, which are more pronounced than the differences at the country levels

9 As long as definitions such as Individualism–Collectivism, Power distance, Uncertainty avoidance, Masculinity–Femininity, and Long term-orientation seem difficult to materialize in specific continuous measures.

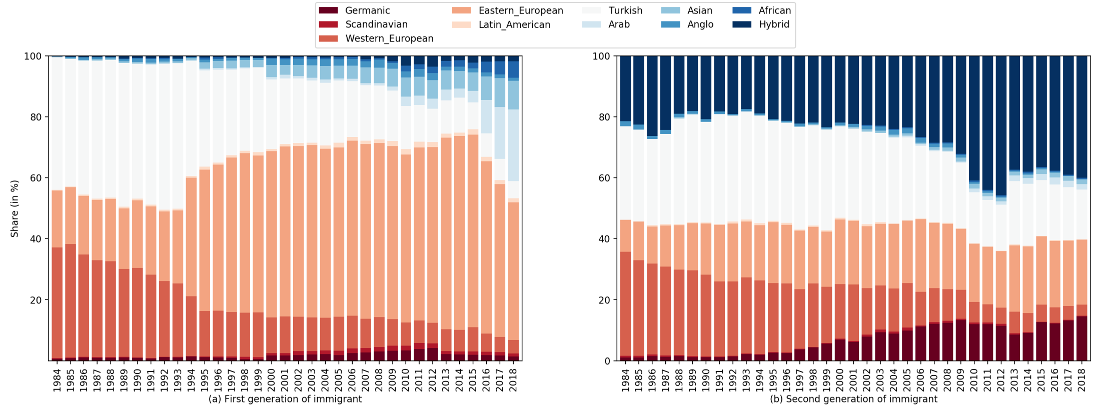


Figure (3) First- and second-generations of immigrants by ethnic group (1984 - 2018)

in SOEP are classified. Therefore resorting to Beugelsdijk et al. (2017),[10] we classify the remaining cultures using the expert input of understudied economies throughout the world reported by Fainshmidt et al. (2018) (see figure 8).[11]

As a result, to identify the cultural origin and the citizenship status of the surveyed individuals, we use the above described cultural classification with two main sources of information, the individual tracking file *ppath* and the parental information survey *bioparen*. Both tables report the country of origin, which is consistent for most individuals, allowing us to reliably classify individuals into cultural clusters.[12] It is important to note, that if both parent's origin is available and each belongs to different cultural clusters, we are not able to assign a culture, as long as we do not know with certainty how work-related values are inherited within a household, therefore labeling it as "hybrid". Figure 3 depicts the proportion of migrants belonging to 10 cultural clusters plus "hybrid" from 1984 to 2018. Note, that the proportion of ethnic German immigrants increased significantly after 1990 and then remained stable from 2009 onward. Also evident is the incoming refugee migration flow occurred in 2015, when the share of Arabs suddenly increased for first-generation immigrants. Finally, the number of surveyed individuals from a given cultural background appears to be correlated with its relative distance to Germany, since our main objective is to compare the immigrant-native gap even after the first generation, due to small sample concerns, we only include Western and Eastern European, and Turkish/Greek cultural clusters in our subsequent analysis.

In addition, determining the origin of the immigrants is key to proxy the implicit costs for transferring skills between two countries suggested by Chiswick's "assimilation effect". Unfortunately, self-assessed German skills are not entirely

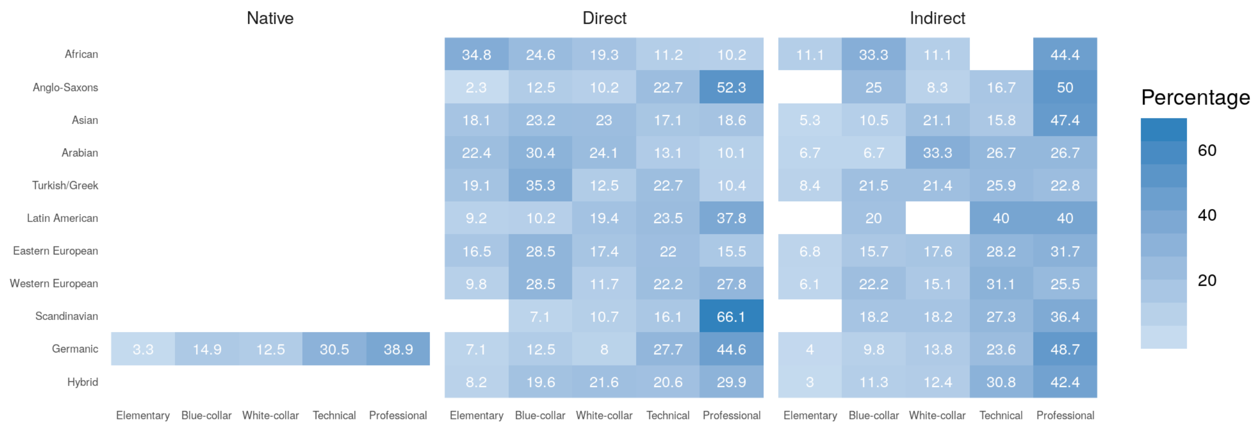
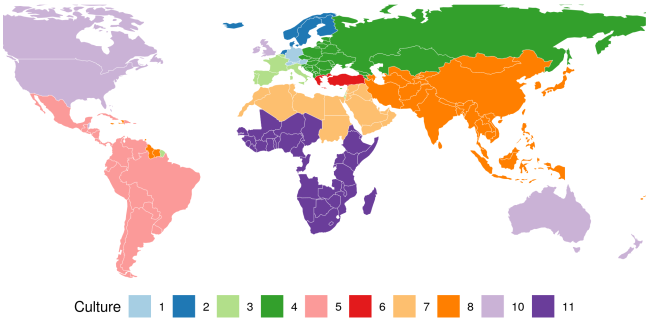


Figure (4) Career choice by ethnic group (aged 20-65)

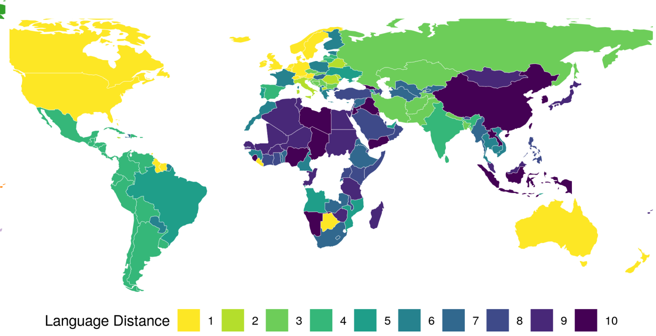
[10] The presence of supra-national cultural zones resonates well with the work on country institutional profiles. (...) [Inasmuch as] institutional environments shape organizational practices and structures and explain their diffusion and spread within and across countries

[11] A side issue could also be raised within our current implementation: Is it reasonable to assume cultural values do not vary over time? Ronen and Shenkar (2013) using historical data finds support for the divergence hypothesis, meaning that despite globalization, culture clusters seem to be more distant over the years. Therefore, our classification could reasonably suit our entire sample range (1984–2018).

[12] Consistency is achieved because the variable retrieves a mix of self-reported data, citizenship information and, family background questionnaires. Nonetheless, we use a series of further filters to certainly determine an individual's cultural background (i) If an indirect migrant does not report any citizenship of their parents nor their own, we drop it from the sample. Then, for the complete sample: (ii) if both parents' origin lay in the same cultural cluster, we classify an individual as belonging to such culture, (iii) if one of the parent's origin is missing, we classify the individual culture according to the available parent's origin. (iii) If no parent's origin is reported, we used the self-reported origin. Lastly, (iv) if the origin of both parents is available, but both belong to different cultural clusters, we classify an individual culture as "hybrid".



(a) Map of cultural clusters.
(1) Germanic (2) Scandinavian (3) Western European (4) Eastern European (5) Latin American (6) Turkish/Greek (7) Arabian (8) Asian (10) Anglo-Saxon (11) African.



(b) Map of Language distance to German:
(1) Closest (10) Farthest.
.

comparable between the different SOEP waves, owing to the fact that the questions and their respective framing have been greatly changed across surveys. In spite of such a drawback, we proxy the country-level language costs regardless of the individual proficiency using the Levenshtein distance with respect to German reported by Isphording and Otten (2011), which in turn seems like a reasonable measure because (i) evaluates the phonetic similarity of two languages in a continuous fashion, (ii) is robust to motivation effects when learning a second-language (iii) its comprehensive,[13] and (iv) appears to be a negative and linear predictor of direct immigrant's individual language-skills for SOEP data (1997-2003) when controlling for efficiency, exposure and economic incentives as demonstrated by Isphording and Otten (2011). To impute the values, in the case of direct immigrants, the minimum value reported between each parent's and own reported country of origin was chosen,[14] Meanwhile, for native and indirect immigrants linguistic distance is set to 0, as we consider that both groups attended a German school and therefore exhibit a fairly high degree of proficiency. Figure 5b shows the country-wise linguistic distance with respect to German grouped in deciles.

3.5. Education and Work Experience

As mentioned in the literature review, education and experience are valued as the main "observable" determinants of cognitive skills accumulation,[15] we must control for such variables in addition to other possible individual "taste" factors such as migration background, gender, culture, and other relevant individual characteristics, as shown in the current section. As shown in figure 1, using the Mincer equation in our model enables us to observe the main distributional characteristics of wages: positive but concave age-earnings profiles, which are also observed in our sample for the upper age-occupation profiles.

It is important to highlight that the aforementioned approach is widely used by the migration literature, on one hand with regard to education, Algan et al. (2010) show that educational investments vary significantly between natives, first- and second-generation immigrants. To demonstrate it, Algan et al. (2010) proxy educational attainment with the age an individual leaves full-time education and uses the country of origin and religion as independent variables. More specifically, they find that all groups of first-generation immigrant men have significantly less education than comparable native German men. For second-generation immigrants the results are less severe; all indirect immigrant groups finished their education at a later age than their first-generation counterparts. Albeit less large, differences between indirect immigrants and native German men still persist. In this paper, however, we can control for this effect using the actual years of education, which is a more robust measure than the *highest obtained degree*, since years of education is a continuous variable, comparable across countries.[16]

13 Few languages were missing from the calculations, and therefore we used the values for the closest linguistic relative, e.g. for Dari (Afghanistan) we use the distance attributed to Persian.

14 This holds as long as we consider that (i) if the information of at least one parent was disclosed, individuals have had some contact and therefore communicate in their parent's mother tongue, (ii) if they possess certain citizenship, at least a basic command of the origin's official language could be expected, and (iii) the associated assimilation costs should be the minimum available in case an individual is proficient in multiple languages.

15 We intentionally focus only on the nowadays traditional cognitive skill approach, however, there is already a vast literature on the determinants and effects of non-cognitive skills in the labor market, e.g. J. J. Heckman and Rubinstein (2001), Cunha et al. (2010) or John and Thomsen (2014) which uses GSOEP data.

16 We do acknowledge the fact that the quality, as well as the content of a year of education across countries, might differ, even exhibiting high levels of heterogeneity between countries e.g. US (state) colleges. Therefore, in the augmented model we added two further dummies reflecting if the individual studied in Eastern Germany or abroad to possibly capture additional skill-transfer costs. Quality/content differences are not directly addressed as long as we do not include *years of education* × *education location* interaction terms.

As described in section 3.6, our observed age range remains quite large (20 years), we still have to control for experience effects. After all, take for example an individual A within the sample that is 35 years old in 2018, and compare it to an individual B that was 35 y/o in 2003, participated in all survey waves of SOEP, and thus was 50 y/o in 2018. For both individuals, the highest-skill occupation registered within 30 and 50 y/o is chosen, but individual B has way more potential work experience and consequently plausibly a highest-skill occupation compared to individual A. Also, as we are including in our sample both men and female, using findings reported by Kelle et al. (2017) and Simonson et al. (2011) which document the existence of marked regional (east-west) and cohort differences on age-full-time and age-part-time experience profiles between women, we define work experience as *full-time experience* + 0.5 × *part-time experience*, since both variables are measured in years, and part-time roughly corresponds to half the exposure of full-time jobs when analyzing the reported weekly hours. Finally, figure 15 shows the observed distribution of years of education, full-time experience, and part-time experience profiles grouped by immigration background.

3.6. Age and years in Germany

As shown in figures 1 and 2, the age-occupation profiles of native Germans between 30 and 60 years old remain stable over time. While there are normal, unproblematic fluctuations in age-occupation profiles of direct and indirect migrants due to a small sample size. This regularity is in line with findings in Mortimer et al. (2002), which state that even when the final decision on what occupation to pursue differs over time and depends on social-economic and labor market conditions, after the 30s, most people have already established themselves in their life-long occupation. In response to this evidence, combined with two further facts; first, most observations are around age 40 ± 10 (see figure 13), and second, we are not able to capture individual heterogeneity in their own occupational choice over time (see section 4.1), we restrict the age range of the sample between 30 and 50 years old instead of choosing one arbitrary age (e.g. 35) and take only the last highest (with respect to skill level) occupation of an individual between the age 30 - 50 as the observed career choice in their whole life-time. This filter allows us to reduce the chance that an individual was observed unemployed due to exogenous business-cycles (see figures 10 and 11 to contrast the variability in year-occupation profiles).

Although it is well documented that duration of stay and age of arrival could be an important predictor of occupational success due to the "assimilation effect". van den Berg and Weynandt (2012) finds that the expected duration of stay is not a good predictor of the actual stay using SOEP data, arguing that individuals systematically underestimate their stay in Germany by using simple heuristics to forecast the future. In addition, duration in Germany could not be defined explicitly for natives and indirect immigrants. In this paper, to achieve some degree of comparability between natives, indirect and direct migrants, we assume that the average effect of one year more in Germany is constant regardless of the age of arrival or cohort.[17] That said, in the augmented model we still control for the possible effect of the number of years in Germany,[18] including its theorized quadratic term, even for those previously defined as 1.5 generation immigrants.

3.7. Cohort effects

According to Lemieux (2006), using cross-sectional data constitute a well-known problem when trying to model the life-cycle choices and earnings profiles, since it can over or understate the actual life-cycle dynamics due to the temporal correlation structure is implicitly neglected or artificially imposed. The use of repeated cross-sections does not improve the panorama, as the year, cohort, and age effects cannot be independently estimated because those variables can be expressed as linear combinations of each other. In that regard, Ma (2020) mentions that the nature of human capital accumulation introduces asymmetric switching costs among individuals across different cohorts, as they solve life-cycle choice problems subject to different market conditions and career prospects, generating high amounts of individual heterogeneity. In the case of immigrants, Carliner (2000) notes that assimilation effects are often confounded with cohort effects, and find negative cohort effects for language abilities for migrants in the US using census data of 1980 and 1990.

In our specific implementation, we cannot model individuals over time, given our method's assumptions (see section 4.1), however, it is worth mentioning that firstly our sample filters are designed in order to remove at some extent the noise produced by fluctuating demand-side conditions across years, enabling us to drop year dummies (see section 3.2). Secondly, in the augmented model we introduce explicitly the assimilation effect faced by direct immigrants, and to some extent also indirect immigrants (those belonging to the 1.5 generation), controlling for the years they have lived in Germany, but as was already mentioned, it assumes that the assimilation is uniform across ages and cohorts, which could be reasonably refuted. Third, to maintain the model parsimony, we refrain from using interaction terms, hence, we assume that odd-ratios and the marginal effects on probabilities of the independent variables remain the same across all observed years, cohorts, and ages, which is a direct consequence of the additive separability property of the utility function used. Fourth, even when we account for cohort effects in the augmented model, we add them just as a level effect on utility as mentioned in the previous point. Finally, to create the cohort dummies, we divide our sample into age terciles, which correspond to the generations *Baby Boom* (1950-1964), *X* (1965-1973) and *Millennial* (1974-1988).

[17] This consideration is also important if we take into account that the cohort, age, and year effects are not identifiable simultaneously in a simple way as they are linear combinations of each other. For further details see 3.7.

[18] Which is equal to 0 for German-born citizens or foreigners.

3.8. Other relevant variables

The migration literature investigates the effect of multiple variables and recognizes the importance of gender, educational quality, marital status, children, parental background, location, duration of stay, and networks. To acknowledge those insights, and when possible, we created variables reflecting such differences, however, we introduce all variables as linear components within the utility function, which implies that instead of having multiplicative (composite) effects or producing an "objective" effect (e.g. differential education quality or contents) the variables act as non-monetary rewards (costs) on utility and should be interpreted accordingly. That said, we are going to briefly present the remaining studied variables, with their respective theoretical or empirical support.

Following Adserà and Ferrer (2014), family formation and occupational choice are arguably interrelated decisions and few studies can provide causal effects for estimates of intermarriage or immigrant fertility adaptation. Even when studies in France, the US, and Canada find that immigrants married to natives have more schooling and earn significantly more than immigrants marrying other immigrants, still exist two possible, mutually exclusive explanations: (i) marriage is driven by unobserved abilities or preferences of individuals or (ii) marriage speeds up the assimilation of the immigrant partner. Therefore, to account for such differences —and without looking for a causal explanation— we added to the augmented model a dummy variable, equal to 1 if the individual is married to a German at the year of the observation.[19] In the same vein, using some insights of Becker (1960), we included a dummy variable equal to 1 if individuals have at least one children at the time of the observation, therefore broadly accounting for the utility change of having children, but recognizing the quantity/quality trade-off, thus not adding it as a continuous variable.[20]

Increasing interest has been raised in migration and occupational choice research to investigate the role and persistence of parental background on children's wage and occupational decisions. For instance, Constant and Zimmermann (2003) notice that parental background has an important weight in determining native Germans's occupational decisions possibly due to an early designation of individual occupational tracks through the staggered access to the schooling system. Conversely, they also observe that for immigrants, the parental background is less persistent and their careers exhibit slight upward social mobility patterns, at least for women. These insights are in line with international studies, as reported by Sweetman and van Ours (2014), who show that, in general, the parental background is also persistent for second-generation immigrants, suggesting a "reset" effect when individuals move between countries. To consider this phenomenon, we used the SOEP reported parental ISCO occupation when respondents were 15 years old, and using the same classifications described in section 3.2 and 3.3, we selected the highest skill level occupation reported between parents.[21] The latter is implemented to simplify the model, although we recognize that mother and father occupations can have differential effects which possibly also depend on the child's sex. For example, if one individual's parent reports working in a *white-collar* job and the other parent in a *blue-collar* job, the *white-collar parent* dummy is set to 1, whereas the other parental occupation dummies are equal to 0.

The distributional differences conditional to regions and sex have also been extensively documented. In the case of east and west Germany differences, Van Hoorn and Maseland (2010) and Smolny and Kirbach (2011) find that occupational distributions and wage differences could not be attributed to values nor to differential patterns of human capital accumulation, but rather a persistent unexplained regional divergence. Nonetheless, as we also observe important differences in our sample, we include in the augmented model a dummy variable reflecting if —at the survey year— an individual lived in any of the former East German states, including Berlin. Furthermore, to control for gender distributional differences and tastes, we also add a *female* = 1 dummy variable in both simple and augmented models. We perceive that, even when some researchers (i.e. Constant and Zimmermann (2003)) also add gender dummies, the high levels of heterogeneity between and within sexes —such as the ones documented in Kelle et al. (2017) and Simonson et al. (2011)— renders the use of dummies and insufficient to reflect hypothesized explanation mechanisms more accurately.

Finally, another variable that is often regarded in the migration literature is migrant networks. Orcutt-Duleep (2014) documents important and negative effects of social networks in occupational success Unfortunately we are not able to observe the location of the individuals at the city or communal level, thus not being able to take such a consideration into account. Additionally, there is no native or indirect German counterpart to those variables with a clear economic or behavioral meaning, forbidding us to directly address them in this study.

[19] In the case of native individuals, the variable could be loosely interpreted as being married, since the proportion of inter-culturally wed Germans is 1 to 4, while the proportion of the non-German wed immigrants is 9 to 1

[20] We conjecture that in practice having children have an ambiguous causal relation: a better-paid occupation leads to more disposable income thus the possibility to have (more) children or more children require more disposable income, hence inducing individuals to obtain a higher paid occupation.

[21] Such a procedure also enable us to increase the sample, since if one parent information was missing, the only available occupation was taken. However, it is important to mention that, as a general rule, the occupation of the mother has higher missing rates, a phenomenon that could be non-random and reveal substantial family composition heterogeneity. Yet the issue requires more exploration.

4. THEORETICAL MODEL

In this section, we discuss the theoretical framework of the estimated model to approximate the effect of differential migration backgrounds on occupational choice. First, we do a brief review of the Multinomial Logit Model (MLM), its main assumptions, and their consequences. Next, we also discuss some differences and trade-offs between the reduced-form and the structural modeling techniques within our current implementation and finish with a quick debate over the implications of the selection bias induced by self-selection in choice-based models.

4.1. The Econometric Model

To frame our econometric approach we track closely the Multinomial Logit Model reported by both Constant and Zimmermann (2003) and Schmidt and Strauss (1975). According to standard economic theory, individuals are assumed to be rational and to have preferences over the complete set of existing occupations. Then, considering an occupational choice model, where each individual can choose between a complete and exhaustive set of $J + 1$ alternatives, indexed as $j = 0, 1, 2, \ldots, J$, the utility of an occupation as a function of individual i characteristics can be approximated by the following linear equations:

$$\begin{aligned} U_i(j = 0) &= \beta_0 X_i + \epsilon_{i,0} \\ U_i(j = 1) &= \beta_1 X_i + \epsilon_{i,1} \\ &\vdots \\ U_i(j = J) &= \beta_J X_i + \epsilon_{i,J}, \end{aligned}$$

Where β_j is a $1 \times p$ vector of parameters, X_i is a $p \times 1$ vector of observed explanatory variables specific for individual i and $\epsilon_{i,j}$ denote the random utility component associated with the specific occupational choice j and individual i. We observe the revealed preference choice j^* if and only if $U(j = j^*) > U(j = k)\, \forall\, j \neq k$, thus, given that individual socio-demographic characteristics do not vary across alternatives, and we don't introduce particular alternative attributes into the model, the choice probability for an alternative j can be expressed as:

$$Prob_{i,j} = Prob(U_j > U_k\, \forall j \neq k) = Prob(\epsilon_j - \epsilon_k < (\beta_j - \beta_k)X_i\, \forall j \neq k)$$

Therefore, the identification of the model relies on the conjectured relation between the density of the errors (random utility) differences and the density of the original errors (random utility),[22] which in the case of the Multinomial Logit is achieved assuming the errors to be independently and identically distributed as a Gumbel or "extreme value" distribution, thus, the choice probability of an alternative j is calculated as:

$$p_{ij} = Prob[j_i^* = j] = \frac{\exp(\beta_j X_i)}{\sum_{k=1}^{J} \exp(\beta_k X_i)}, \quad j = 1, \ldots, J. \text{ Normalized to the baseline alternative } j = 0.$$

For any two alternatives j and k, the normalized log-odds ratio of the logit probabilities is then:

$$ln\left[\frac{p_{ij}}{p_{ik}}\right] = (\beta_j - \beta_k)x_i$$

Note that this ratio does not depend on any alternatives other than j and k. That is, the relative odds of choosing j over k are the same no matter what other alternatives are available or what the attributes of the other alternatives are. This is known as the Independence of Irrelevant Alternatives (IIA), which is a restrictive assumption arising naturally from the error terms' distributional assumptions. Nonetheless, McFadden (1973) suggests that in cases where the outcome categories "can plausibly be assumed to be distinct and weighed independently in the eyes of each decision maker," a Multinomial Logit Model can be reasonable used.[23]

In our favor, two characteristics of the IIA support our current implementation. First, from the very definition of the skill groups, we clearly differentiate the nature of the work performed —its sector or language intensity—, the level of formal education, and the amount of informal on-the-job training, which are objective factors that could be also reasonable regarded by decision makers, thus validating the use of the model in broad terms. Second, as Train (2009) argues, IIA ensures that the exclusion of alternatives in estimation does not affect the consistency of the estimator, and since we are only interested in examining occupational choices, and not for example the decision between joining the labor force or not, unemployed individuals can be excluded from the analysis without severe consequences.

22 However, there is an infinite number of densities for the J error terms that give the same density for the $J - 1$ error differences, hence, to solve the model, everything has to be normalized with respect to one alternative. Such a transformation is consistent for random utility models because "only relative utility matters" (Train, 2009).

23 There are also some statistical tests to validate the IIA assumption, however, according to Train (2009), more than assessing the statistical compliance with IIA, these tests do not provide as much guidance on the correct specification to use instead of Logit, therefore be used only as a measure of "appropriateness". The construction of the tests, as well as their results for the estimated models are reported in the Appendix.

In addition to the odd-ratios—and to enhance the model's interpretability—we report the Average Marginal Effects (AMEs) of the relevant independent variables. Basically, the Marginal Effects are calculated as the derivatives of the choice probabilities with respect to a particular independent variable z, defined as:

$$\frac{\partial P_{i,j}}{\partial z_{i,j}} = \frac{\partial \frac{\exp(\beta_j X_i)}{\sum_j \exp(\beta_j X_i)}}{\partial z_{i,j}} = \frac{\partial B_j X_i}{\partial z_{i,j}} P_{i,j}(1 - P_{i,j}) = \beta_z P_{i,j}(1 - P_{i,j}). \text{ Since } B_j X_i \text{ is linear in } z_i \text{ with coefficient } \beta_z$$

which could be also stated in discrete form (see Long and Freese (2014) for further details). Then, predicted probability values given a change in z, *ceteris paribus*, are averaged across individuals and the AME with respect to z is recovered. However, since the derivatives must not be equal at all values of a continuous z, we also report the derivatives conditional on the observed individual characteristics in the form of plots. The main advantage of this approach is that it allows the direct interpretability of the coefficients and helps us to compare across models.

Moreover, weighing the limitations of the Multinomial Logit Model allows a proper selection of the design together with a fair judgment of the obtained results, from which Train (2009) mentions three: (i) the logit model implies proportional substitution across alternatives, given the researcher's specification of representative utility, (ii) logit can model systematic taste variation (that is, taste variation that relates to observed characteristics of the decision maker) but not random taste variation, and (iii) logit cannot handle situations where unobserved factors are correlated over time. The first limitation is another way to describe the IIA, which in turn was already assessed above. The second limitation, more than an obstacle, circumscribe the scope of the answer to our research question, which wants to appraise the systematic differences between immigrants and native Germans when choosing an occupation, however, by the nature of the model we are not able to disentangle if these observed choices are driven solely by tastes or also by social factors such as discrimination, since, even if the latter is systematic, it is not directly observed by us, in consequence, both effects are captured simultaneously by the error term $\epsilon_{i,j}$.[24] Finally, modeling occupational inter-temporal decisions with a logit model is not appropriate because it imposes strong constraints on the temporal correlation structure, therefore, to maintain simplicity we pooled our sample to a cross-sectional shape, carefully selecting the last available observation of the surveyed individuals between 30 and 50 years old, where the occupational choices seem to reach a stable distribution, hence, retaining only plausibly time-independent observations.[25][26]

4.2. Alternative Structural Models

Non-structural estimation is concerned with recovering the relationship of key parameters of interest using exogenous in-sample variation without the restriction of causality in the relationship. The failure of reduced-form modeling lies in failing to interpret the correlations shown in the model as causation (reverse causation or outside factors) or more specifically that those estimated relations are not policy invariant, thus, reflecting just *ad hoc* statistical relationships (Chetty, 2009). For example, in the analysis of occupational choices, the fundamental reason why using an extension of the Mincer model could yield a biased estimation is that schooling and occupational choices are not randomly assigned, so the higher-rank occupation (highly-educated people) and lower rank occupation (low-educated people) may systematically differ from each other, both from observed and unobserved characteristics.

In contrast, the structural estimation is concerned with recovering some or all the structural (primitive) elements of the model at the expense of assuming specific functional forms reliant on theoretical informed assumptions, independent of the particular sample(s). Moreover, to achieve internal consistency, Structural modeling often needs "conditioning variables" that are not explicitly part of the economic theory as a way of controlling for plausible differences across observations and securing identifiability (Chetty, 2009). That said, next, we briefly recapitulate Keane and Wolpin (2009) explanation of the most general form of a structural model in occupational choice problems, where the individual's decision rule at each age is given by:

$$d_{mt} = \begin{cases} 1 & \text{if } (\widetilde{D}_{mt}, \widetilde{X}_{mt}, \tilde{\epsilon}_{it}) \geq 0 \\ 0 & \text{if } (\widetilde{D}_{mt}, \widetilde{X}_{mt}, \tilde{\epsilon}_{mt}) \leq 0 \end{cases}$$

$\widetilde{D}_{mt}$ is vector of the history past choices ($d_{mt} : t = 1, \dots, t-1$), $\widetilde{X}_{it}$ is a vector of contemporaneous and lagged values of J additional variables ($X_{mjt} : j = 1, \dots, J; t = 1, ..., t$) that enter the decision problem, and $\tilde{\epsilon}_{it}(\epsilon_{mt} : t = 1, ..., t-1)$ is a vector of contemporaneous and lagged unobservables that also enter the decision problem. Now let $d_m(a) = 1$ if alternatives m

24 In respect of such a discussion, Arrow (1998) makes a thorough analysis on how the economic science's approach cannot, in general, untwine both utilitarian and discriminatory determinants of choice on account of the impossibility to measure them appropriately.

25 Although we still keep observations from individuals sharing the same household, thus, to correct such a correlation, all our results used clustered errors at the household level.

26 We didn't analyze if there is systematic attrition in our sample, nor took into account the oversampling of migrants in comparison to the general population. Two important concerns questioning the randomness of our sample and hence the external validity of our results. See some deeper thoughts on the topic in subsection 4.3.

is chosen ($m = 0, 1, \dots, J$) at age a and zero otherwise. The reward per period at any age a is given by:

$$R(a) = \sum_{k=1}^{J} R_m(a) d_m(a)$$

where $R_m(a)$ is the reward per period associated with mth alternative. At any age, the individual's objective is to maximize the expected present value of the remaining lifetime reward. $V(S(a), a)$ be the maximum expected present discounted value of remaining lifetime utility at age a and given the state space $S(a)$ and discount factor δ

$$V(S(a), a) = \underset{d_m(a)}{maxE} \left[\sum_{t=a}^{A} \delta^{t-a} \sum_{m=0}^{J} R_m(a) d_m(a) | S(a) \right]$$

A structural model can be useful in occupational choice problems as the model can directly account for heterogeneity in an individual's characteristics and preferences. Furthermore, if correctly specified, Structural models also help to provide more straightforward and economically meaningful parameters, which in turn play an important role in simulating counterfactuals and calculating diverse arrays of policy effects on agents. However, it is worth mentioning that one of the drawbacks of non-reduced modeling is that data and computational burdens may render them impractical in very complex environments such as the ones depicted in an occupational choice problem.

4.3. Selection bias and self-selection

Finally, an important issue for many economic models —especially for those of occupational choice— is the selection bias driven by self-selection. In words of J. J. Heckman (2010): "[Such a] problem (...) arises when a rule other than simple random sampling is used to sample the underlying population that is the object of interest", which in this case is determined by the agent's observed decisions j^*. In the simpler case, self-selection introduces a bias that undermines the external validity of the model, as long as the estimated coefficients implicitly include the mean of the sampling error conditional distribution—which renders them biased and inconsistent—. In the case of choice based sampling, as J. J. Heckman (2010) demonstrates, the conditional on covariates sample density of alternatives $g(j^*|X^*)$ is equal to:

$$g(j^*|X^*) = f(j^*|X^*) \times \left\{ \left[\frac{\omega(j^*)}{\sum_{i=1}^{I} \omega(i) f(i)} \right] \times \left[\frac{1}{\sum_{i=1}^{I} f(i|X^*) \frac{g(i)}{f(i)}} \right] \right\}$$

Where $f(j^*|X^*)$ is the population conditional on covariates density of alternatives, $\omega(.)$ is a non-negative sample weighting function, and i indexes individual types over the full population $i = 1, 2, \dots, I$, therefore, highlighting explicitly how the sampling rule deviates the sample proportions from the population proportions. Although Manski and McFadden (1981) directly address such a sampling problem, we are unfortunately unable to make *a priori* assumptions of β_J^* to determine the optimal sample design and offer logit's Maximum Likelihood Estimators (MLE) robust to self-selection, oversampling and non-random attrition.[27]

5. RESULTS

In the following subsections, we first give an interpretation of the coefficients for our models. Next, the results and economic interpretations are presented. Since the results of the ISCO-skill and language-skill models are similar, we will review them in the same subsection. The results for all the specifications can be found in the appendix; for the two dependent variables ISCO-skill and Language-skill, a basic and augmented model, as well as their respective log-odds and average marginal effects. The augmented model is the extended version of the basic one with additional independent variables. The accuracy, however, does not seem to improve with the inclusion of more variables for both tasks. Finally, we modify the basic and augmented models to exclude either the variable language distance (e.g. *Model 1* in table B), or the variables direct and indirect migrants (*Model 2*), because it appears that language distance and migration background are capturing the same phenomenon, and when included at the same time, migration background loses its significance —especially the indirect migration dummy—, probably due to the fact that it is capturing less heterogeneity compared to the language distance counterpart. In general, since we are interested in modeling Native-Immigrant Gaps, we chose those models with migration backgrounds instead of language distances, however, the language distance remains an interesting variable to further explorations.

[27] Moreover, since such a phenomenon affects the underlying assumptions of the MLE—knowing both the marginal and population distributions to consistently estimate the covariance structure—, alternative modified versions of MLE should be implemented to correctly estimate the model. However, it is also important to note that such deviations fit more naturally in a pure structural model, e.g. J. J. Heckman (1976), thus, rendering our analysis as a first, preliminary reduced-form approximation to an otherwise more appropriate, fully-fledged occupational choice model.

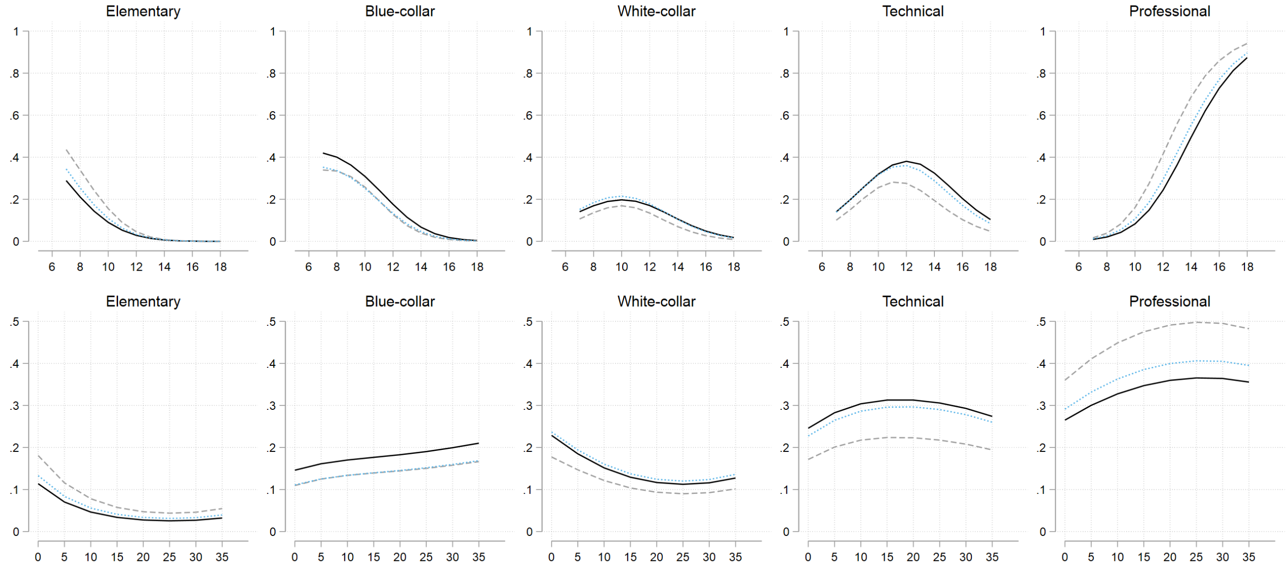


Figure (6) Average Marginal Effects of the continuous variables estimated in the basic model for ISCO Skill discriminated by migrant status. (Top) Years of education. (Bottom) Years of experience. In addition, Solid = Native, Dotted = Indirect Migrant, Dashed = Direct Migrant

5.1. Interpretation of Coefficients

Instead of directly showing the log-odd ratio estimates, we present the tables which show the marginal effects in order to provide a clearer overview of the relationship between explanatory and dependent variables. The coefficients indicate the change in possibility and likelihood that a person would choose a specific occupation given that there is 1 unit increase in the corresponding independent variables. For education and working experiences, 1 unit increase corresponds to one more year in school or in the job market. On the other hand, for the dummy variables, such as female, 1 unit increase simply implies that our observation is a woman (0, otherwise). If we for example take table 7, the probability to be working in the fourth quantile of language skill if you are from Eastern European origin is -8.72%, significant at the 1% level. In other tables, for example, table 11, the log-odds instead of the average marginal effects are given. The log-odds have to be interpreted by taking the exponential to give the odds ratio. For example, for model (1) in table 11, the *odd* of having a white-collar occupation as a direct migrant is $e^{-0.705} = 0.4941$, meaning that a direct migrant has a $1 - 0.4941 = 0.5059$ or 50.59% less likely chance of working in a white-collar occupation compared to a native German.

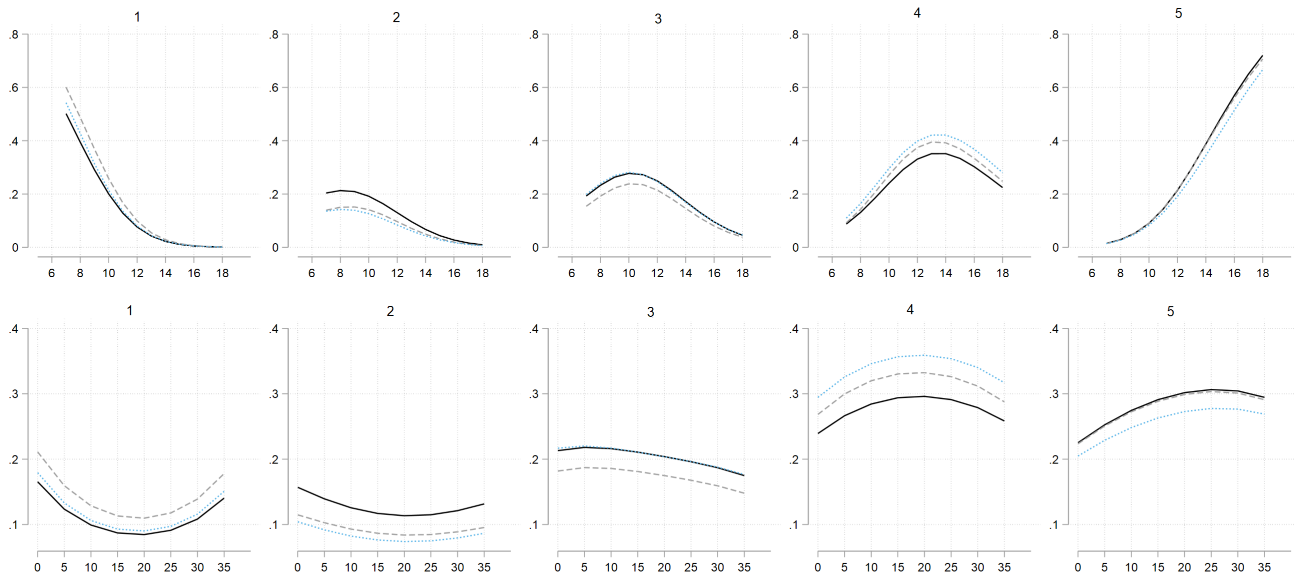


Figure (7) Average Marginal Effects of the continuous variables estimated in the basic model for Language Interactivity Skill discriminated by migrant status. (Top) Years of education. (Bottom) Years of experience. In addition, Solid = Native, Dotted = Indirect Migrant, Dashed = Direct Migrant

5.2. Economic Implications

In figure 6 and 7 the average marginal effects are plotted for both dependent variables and with years of education and experience as independent variables. Note the scaling of the figures is different across rows.

Our hypothesis regarding language-interactivity skill was, in line with the literature review, that natives transition to higher language-intense jobs, while immigrants obtain lower language-intense jobs. In our research, this only partly holds. In figure 7 compared to the average marginal effects with respect to ISCO skill, it is shown that native Germans have advantages over immigrants in language skill groups 2 and 3. For groups 1 and 4, direct immigrants have a higher probability given the years of education or experience. Remarkable is the highest language skill group. Here, we would expect the largest differences between natives and direct immigrants but the opposite is shown. The results are very similar to the ISCO language skill (figure 6), which can be explained by the fact that most occupations have the same position with respect to both ISCO skills and language skills (table D). Therefore, we will discuss both dependent variables in this section together.

The high educational level remains favorable for professionals and for the highest language skill; a positive and significant marginal effect of years of education is observed (table 7 and 3). This result supports the classic human capital theory in the sense that a higher educational background encourages people to pursue a more skill-demanding occupation. The difference between natives (solid) and direct migrants (dashed) for the language skill is almost non-existent in the highest language skill quantile. We deem this to be related to a translation error resulting from the O*NET to SOEP conversion. Take for example ISCO code 2529 in table D; data mining analyst. This occupation is in the "professional" category and is also in the highest language interactivity skill quantile. While we agree one needs a high knowledge of the language for this specific occupation, it does not need to be German per se. This is one of the examples where a translation of the "need for good English skills" to "need for good Germans skills" is most likely false. From figure 6, we observe that the marginal effect of education is higher for direct migrants in the professional sector, whereas the effect is almost identical for native Germans and indirect migrants. This could imply that direct immigrants, in some respects, possess several advantages over Germans. One major advantage might be better skills in other foreign languages other than German, for instance English, which has become almost mandatory in managerial classes and academic circles. Other advantages such as working experience abroad or obtaining a degree abroad[28] could as well be helpful for a professional occupation. In comparison with the upward sloping marginal effect for the professional sector, a generally decreasing marginal effect and a hump-shaped pattern are spotted for other sectors for both dependent variables. This may imply that instead of schooling, vocational and related skill training programs could be more crucial in these industrial categories. This effect might be even more obvious in Germany, in which the dual and vocational training system has been long adopted. In addition, an optimal number of years of education regarding different occupational sectors could be potentially inferred from the figure. For elementary and blue-collar jobs, more years of education have a negative impact on the possibility, whereas the supportive years of education for white-collar and technical sectors are around 10 and 12 years, respectively.

The coefficients of working experience are found to be similarly significant but with a limited effect, either positive or negative, in four sectors except technical. The effect of working experience on financial returns is undeniable and has been proven by numerous papers. Yet, its impact on occupational choice could not be so easily concluded.[29] Certainly, for some professional jobs, more working experience could imply future promotion, but in general, the relation between working experience and the occupational choice is obscure.

A distinct significant effect of a person's cultural background on their occupational choice or language skill is not spotted in our results. We do notice that for Eastern European, Turkish and Greek immigrants, the possibility for them to have an occupation in the fourth language quantile is lower. In the augmented model with ISCO skill as the dependent variable (table 6), we see that Turkish/Greek have a significant negative chance of working in the professional sector. A possible reason could be the enforcement of the guest workers program by West Germany in the 1960s. Because of the labor agreement, many immigrants from Eastern Europe and Turkey moved to Germany and worked in either the elementary or blue-collar sector. As pointed out in Constant and Zimmermann (2003) that there exists evidence that individuals inherit their social status and their position in the occupational distribution, we might be able to infer that the negative coefficients could be caused by the disadvantage of lacking social connections or endowments instead of the cultural background. For the dataset consists of solely second-generation immigrants, country of origin seems to have no impact on a person's occupational choice. All variables regarding the origin of a person are not statistically significant in all five sectors. We may be able to say that in Germany, working opportunities could be relatively equal and fair for people with different cultural backgrounds, given they had the same schooling. Another interpretation could also be that the cultural background, to some extent, does not determine or deter a person from doing a specific job. Instead, it could be other family background factors, for instance, income or occupation of parents which make the difference in a person's career life.

[28] The possibility for a person to engage in the professional sector with their education received abroad is 4.31% and significant under 0.05 according to Table 6.

[29] For instance, it is hard to say that a person with working experience in housekeeping would suddenly become a professional scholar.

There exist certain signs of persistence in occupational choices. Some specific advantages are found on those children whose parents work in the Technical and Professional sectors. Leppel et al. (2001) suggests that both male and female students whose fathers are in professional or executive occupations were more likely to choose to major in engineering and the sciences. Hence, their future engagement in technical and scientific fields is expected. In addition, considering that people with higher educational levels would be more likely to teach or have more resources to invest in their children, this finding seems to be relatively intuitive. This argument is further supported by Chevalier et al. (2013), in which both parental education and income have positive and significant impacts on the post-compulsory education of their children. Furthermore, if their parents are already working in the sector, it could be more likely that children would have more social connections and therefore are able to make use of these relationships in the future while looking for a job. We would like to mention also that given their parents do not work in the elementary sector, the possibility for children to engage in the elementary occupation is negative and significant for all other four categories (Blue- and White-Collar, Technical, and Professional). This could be potentially explained by the improvement in occupational skills. In line with the other results so far, for the language skills of parents a similar trend is found. If the highest language skill of (one of) the parents is the second quantile, the child has a positive and significant effect of also having an occupation in the second quantile.

Migrants and Germans are different in their characteristics. However, since we control for many observables, the significance of the actual variables of Direct Migrant and Indirect Migrant are almost not significant. What we do see is that direct immigrants have a 12% higher possibility to engage in professional jobs than native Germans. Considering the fact that Germany is one of the leading research centers across fields, there might be scholars and researchers who come to Germany in search of working opportunities. According to the annual report by Statistisches Bundesamt, the percentage of immigrants to Germany between age 25 to 35 with higher education degree substantially increases from 16.7% in 2005 to 27% in 2017, in comparison to the decrease in the number of immigrants without previous vocational training.[30] Additionally, pursuing a higher education degree in Germany gradually becomes popular among international students in recent decades due to its high quality and relatively low cost; some of those who obtain their degree in Germany might eventually decide to stay and look for a job here. This could be a potential reason why being a direct immigrant actually has a higher possibility to be in the professional sector. With respect to language skills, we already expected that being an indirect migrant should have little effect on the quantile since indirect migrants are born in Germany and learned German since they were young and thus should be on the same level as native Germans.

Women's Engagement in Technical and White-Collar Sectors is positive and significant. Constant and Zimmermann (2003) provides an insight that sex significantly affects occupational choices even after we control for human capital and other characteristics, with women being sorted into white or professional jobs. These findings can further be supported by statistics from the Statistisches Bundesamt (2020). We observe that the female labor force dominates in several sectors. In Human Health and Social Work Activities (WZ08-Q) and Other service activities (WZ08-S), [31] for instance, the ratio of men-to-women is around 1:3 and 1:2, respectively. In Financial and Insurance Activities (WZ08-K), and Professional, Scientific and Technical Activities (WZ08-M), we can see that the number of men and women engaging in these two industry sectors is almost identical. On the contrary, Agriculture, Forestry and Fishing (WZ08-A), Manufacturing (WZ08-C), and Construction (WZ08-F) continue to be male-dominated sectors. Some might notice that in the "Elementary " sector, the coefficient for female is also positive. We would like to mention that instead of some traditional works such as Agriculture, Forestry and Fishing and Mining, our classification includes also Housekeeping and Cleaning which makes up a large proportion of our samples. Therefore, from the table, we would detect a positive coefficient for female also in the elementary sector. Along with these statistics, we might also be able to say that even though there still exist some traditionally regarded male- or female-specific jobs, the working opportunities in some professional areas for both genders seem to be equal and fair. We can not tell whether there are differences in earnings for men and women while doing the same job since that is not in our scope. What is remarkable is that women have a negative and significant chance of working in language quantiles 1 and 2, and positive and significant in quantiles 3, 4 and 5 (table 9). This could be explained by the fact that men tend to do more labor-intensive jobs and women more "white-collar" jobs, where language skills are more important; a comparable example is that STEM majors have a high male/female ratio compared to humanities where the male/female ratio is rather low (Watt, 2010).

6. CONCLUSION

In this paper, we presented an analysis of the occupational status gap between native Germans and immigrants in the influences of human capital variables and cultural background characteristics. We reviewed literature in human capital theory in the context of immigrants and provided descriptive evidence and some analyzes on trends and characteristics of the 1984-2018 period immigrant sample in Germany. We focused our research on two questions: (i) are there significant occupational gaps between immigrants and native Germans and (ii) do immigrants choose less language-intensive

30 Page 80-81, Migration und Integration, Integrationsindikatoren, Statistisches Bundesamt (Destatis), 2019

31 Nursing and midwifery associate professionals is classified under Technical according to ISCO.

occupations.

Using a skill-cell approach (job skills and language skills) and introducing controls for education and experience through the Mincer equation, we found a very small and insignificant effect of ethnicity on career choices of immigrants. However, the results show that gender differences continue to be a prominent phenomenon in the labor market. There are many issues related to classification and measurement errors that we have not solved completely. Focusing on skill concepts, we ignore the unemployment outcome as it has no implication of job skills or language skills. This could be a problem when taking into account that the unemployment rate of first-generation immigrants could be significantly higher than of second-generation immigrants. As a result, we could not observe the progress for labor market performance between these two generations of immigrant.

Furthermore, other problems related to the economic integration of migrants need to be considered but remain largely unaddressed in this paper. In our basic model, using the Mincer equation as the core of our model, it is assumed that the initial starting skills, return to schooling and experience (the slope coefficients) are identical across individuals. We did not capture heterogeneous effects over individuals. Then, we took the observed occupation with the highest skill level of an individual in the age range 30-50, we assume that all observed effects of the first generation of immigrants are the products of pre-migrant human capital and ignore the improvement in one's assimilation process. However, when considering the economic theory of assimilation in the augmented model, the controlled variables for the heterogeneity of the migration and assimilation processes between and within ethnic groups (i.e., years-since-migration profile, the region of residence, cohort effects, the presence of children, education in home country) enter the function linearly and without interaction terms. This makes interpretation of the parameters problematic when the observed effects could not be interpreted as a specific effect for any group of immigrants but as an average effect for all groups. In addition, besides the differences in individuals's heterogeneity, there also exist other external heterogeneity factors at the aggregate-group level should be considered. To be more specific, the proportion of immigrants in one's area, which occupation they have, or what one's ethnic group average-skill level is, also have a substantial effect on the human capital accumulation process and occupational choices of each immigrant in these communities.

Since ethnic labor-market inequalities and inter-generational assimilation are intertwined topics, this paper still exists shortcomings, the answer to these would benefit the field of labor economics. This paper also raises two questions on future research of occupational gaps: (i) what is the role of institutions and policies affecting the immigrants's integration process in terms of labor market discrimination on gender and ethnic identity? And, (ii) how effectively could the structural model capture the heterogeneity of immigrants?

7. REFERENCES

Adserà, A., & Ferrer, A. (2014). Immigrants and demography: Marriage, divorce, and fertility. In B. Chiswick & P. W. Miller (Eds.), *Handbook of the economics of international migration* (pp. 315–367). Elsevier.

Algan, Y., Dustmann, C., Glitz, A., & Manning, A. (2010). The economic situation of first and second-generation immigrants in france, germany and the united kingdom. *The Economic Journal*, *542*(120), F4–F30.

Amuedo-Dorantes, C., & De La Rica, S. (2011). Complements or substitutes? task specialization by gender and nativity in spain. *Labour Economics*, *18*(5), 697–707.

Arrow, K. J. (1998). What has economics to say about racial discrimination? *Journal of economic perspectives*, *12*(2), 91–100.

Baert, S., Heiland, F., & Korenman, S. (2016). Native-immigrant gaps in educational and school-to-work transitions in the 2nd generation: The role of gender and ethnicity. *De Economist*, *164*. https://doi.org/10.1007/s10645-016-9273-4

Basilio, L., Bauer, T., & Kramer, A. (2017). Transferability of human capital and immigrant assimilation: An analysis for germany. *LABOUR*, *31*. https://doi.org/10.1111/labr.12096

Bauer, T., Dietz, B., Zimmermann, K. F., & Zwintz, E. (2005). German migration: Development, assimilation, and labour market effects. *European Migration: What do we know*, 197–261.

Becker, G. S. (1960). An economic analysis of fertility. *Demographic and economic change in developed countries* (pp. 209–240). Columbia University Press.

Beugelsdijk, S., Kostova, T., & Roth, K. (2017). An overview of hofstede-inspired country-level culture research in international business since 2006. *Journal of International Business Studies*, *48*(1), 30–47.

Bodvarsson, Ö. B., Simpson, N. B., & Sparber, C. (2014). Migration theory. In B. Chiswick & P. W. Miller (Eds.), *Handbook of the economics of international migration* (pp. 3–51). Elsevier.

Brücker, H., Kosyakova, Y., & Vallizadeh, E. (2020). Has there been a "refugee crisis"? new insights on the recent refugee arrivals in germany and their integration prospects. *SozW Soziale Welt*, *71*(1-2), 24–53.

Carliner, G. (2000). The language ability of us immigrants: Assimilation and cohort effects. *International Migration Review*, *34*(1), 158–182.

Cheng, S., & Long, J. S. (2007). Testing for iia in the multinomial logit model. *Sociological methods & research*, *35*(4), 583–600.

Chetty, R. (2009). Sufficient statistics for welfare analysis: A bridge between structural and reduced-form methods. *Annu. Rev. Econ.*, *1*(1), 451–488.

Chevalier, A., Harmon, C., O' Sullivan, V., & Walker, I. (2013). The impact of parental income and education on the schooling of their children. *IZA Journal of Labor Economics*, (8).

Chiswick, B. R. (1978). The effect of americanization on the earnings of foreign-born men. *Journal of political Economy*, *86*(5), 897–921.

Chiswick, B. R. (1979). The economic progress of immigrants: Some apparently universal patterns. *Contemporary economic problems*, *1979*, 357–399.

Chiswick, B. R., & Miller, P. W. (2010). Occupational language requirements and the value of english in the us labor market. *Journal of Population Economics*, *23*(1), 353–372.

Constant, A., & Zimmermann, K. F. (2003). Occupational choice across generations. *IZA*, (975), 18–19.

Cunha, F., & Heckman, J. (2007). The technology of skill formation. *American Economic Review*, *97*(2), 31–47. https://doi.org/10.1257/aer.97.2.31

Cunha, F., Heckman, J. J., & Schennach, S. M. (2010). Estimating the technology of cognitive and noncognitive skill formation. *Econometrica*, *78*(3), 883–931.

Diehl, C., & Granato, N. (2018). *Germany : Intergenerational inequalities in the education system and the labour market for native-born children of immigrants from turkey and the former yugoslavia.*

Dustmann, C. (2004). Parental background, secondary school track choice, and wages. *Oxford Economic Papers*, *56*(2), 209–230.

Fainshmidt, S., Judge, W. Q., Aguilera, R. V., & Smith, A. (2018). Varieties of institutional systems: A contextual taxonomy of understudied countries. *Journal of World Business*, *53*(3), 307–322.

Fellini, I., Guetto, R., & Reyneri, E. (2018). Poor returns to origin-country education for non-western immigrants in italy: An analysis of occupational status on arrival and mobility. *Social Inclusion*, *6*, 34. https://doi.org/10.17645/si.v6i3.1442

Fleischhauer, K.-J. (2007). A review of human capital theory: Microeconomics. *University of St. Gallen, Department of Economics Discussion Paper*, (2007-01).

Freese, J., & Long, J. S. (2000). Tests for the multinomial logit model. *Stata Technical Bulletin*, *10*(58), 1–10.

Hartmann, J. (2014). Do second-generation turkish migrants in germany assimilate into the middle class? *Ethnicities*, *16*. https://doi.org/10.1177/1468796814548234

Hartmann, J. (2016). Assimilation over the life course? the career mobility of second-generation turkish men in germany. *Zeitschrift für Soziologie*, *45*. https://doi.org/10.1515/zfsoz-2015-1016

Heckman, J. J. (1976). A life-cycle model of earnings, learning, and consumption. *Journal of political economy*, *84*(4, Part 2), S9–S44.

Heckman, J. J. (2010). Selection bias and self-selection. *Microeconometrics* (pp. 242–266). Springer.

Heckman, J. J., & Rubinstein, Y. (2001). The importance of noncognitive skills: Lessons from the ged testing program. *American Economic Review*, *91*(2), 145–149.

Hess, C., & Green, S. (2016). Introduction: The changing politics and policies of migration in germany. *German Politics*, *25*(3), 315–328.

International Labour Office. (2012). *International standard classification of occupations: Structure, group definitions and correspondence tables* (1st ed.).

Isphording, I. E., & Otten, S. (2011). Linguistic distance and the language fluency of immigrants. *Ruhr Economic Paper*, (274).

Jain, A., & Sabirianova Peter, K. (2017). *A joint hazard-longitudinal model of the timing of migration, immigrant quality, and labor market assimilation* (tech. rep.). IZA Discussion Papers.

John, K., & Thomsen, S. L. (2014). Heterogeneous returns to personality: The role of occupational choice. *Empirical Economics*, *47*(2), 553–592.

Keane, M. P., & Wolpin, K. I. (2009). Empirical Applications of Discrete Choice Dynamic Programming Models. *Review of Economic Dynamics*, *12*(1), 1–22. https://doi.org/10.1016/j.red.2008.07.001

Kelle, N., Simonson, J., & Gordo, L. R. (2017). Is part-time employment after childbirth a stepping-stone into full-time work? a cohort study for east and west germany. *Feminist Economics*, *23*(4), 201–224.

Kvetan, V. (2014). Coping with changes in international classifications of sectors and occupations: Application in skills forecasting. research paper no 43. *Cedefop - European Centre for the Development of Vocational Training*.

Lemieux, T. (2006). The "mincer equation" thirty years after schooling, experience, and earnings. *Jacob mincer a pioneer of modern labor economics* (pp. 127–145). Springer.

Leppel, K., Williams, M. L., & Waldauer, C. (2001). The impact of parental occupation and socioeconomic status on choice of college major. *Journal of Family and Economic Issues*, 373–393.

Long, J. S., & Freese, J. (2014). *Regression models for categorical dependent variables using stata* (3rd ed.). Stata press.

Ma, J. (2020). High skilled immigration and the market for skilled labor: The role of occupational choice. *Labour Economics*, *63*, 101791.

Manski, C. F., & McFadden, D. (1981). Alternative estimators and sample designs for discrete choice analysis. *Structural analysis of discrete data with econometric applications*, *2*, 2–50.

McFadden, D. (1973). Conditional logit analysis of qualitative choice behavior. *Frontiers in Econometrics*, 105–142.

Mincer, J. (1958). Investment in human capital and personal income distribution. *Journal of political economy*, *66*(4), 281–302.

Mincer, J. (1974). *Schooling, experience, and earnings*. National Bureau of Economic Research, Inc.

Miyar-Busto, M., Mato Diaz, F., & Gutierrez, R. (2020). Immigrants' educational credentials leading to employment outcomes: The role played by language skills. *Revista Internacional de Organizaciones*, 167–191. https://doi.org/10.17345/rio23.167-191

Mortimer, J. T., Zimmer-Gembeck, M. J., Holmes, M., & Shanahan, M. J. (2002). The process of occupational decision making: Patterns during the transition to adulthood. *Journal of Vocational Behavior*, *61*(3), 439–465.

Orcutt-Duleep, H. (2014). The adjustment of immigrants in the labor market. In B. Chiswick & P. W. Miller (Eds.), *Handbook of the economics of international migration* (pp. 108–171). Elsevier.

Peri, G., & Sparber, C. (2009). Task specialization, immigration, and wages. *American Economic Journal: Applied Economics*, *1*(3), 135–69.

Ronen, S., & Shenkar, O. (2013). Mapping world cultures: Cluster formation, sources and implications. *Journal of International Business Studies*, *44*(9), 867–897.

Schmidt, P., & Strauss, R. P. (1975). The prediction of occupation using multiple logit models. *International Economic Review*, 471–486.

Schüller, S. (2012). Parental ethnic identity and educational attainment of second-generation immigrants. (443). https://ideas.repec.org/p/diw/diwsop/diw_sp443.html

Sebastian, R., & Ulceluse, M. (2019). The effect of immigration on natives' task specialisation: The case of germany. *International Journal of Manpower*.

Simonson, J., Gordo, L. R., & Titova, N. (2011). Changing employment patterns of women in germany: How do baby boomers differ from older cohorts? a comparison using sequence analysis. *Advances in Life Course Research*, *16*(2), 65–82.

Smolny, W., & Kirbach, M. (2011). Wage differentials between east and west germany: Are they related to the location or to the people? *Applied Economics Letters*, *18*(9), 873–879.

Statistisches Bundesamt. (2020). Bevölkerung und erwerbstätigkeit. bevölkerung mit migrationshintergrund-ergebnisse des mikrozensus 2019 [population and employment. population with migrant background-results of the microcensus 2019].

Sürig, I., & Wilmes, M. (2015). *The integration of the second generation in germany: Results of the ties survey on the descendants of turkish and yugoslavian migrants*. https://doi.org/10.1515/9789048526970

Sweetland, S. R. (1996). Human capital theory: Foundations of a field of inquiry. *Review of educational research*, *66*(3), 341–359.

Sweetman, A., & van Ours, J. C. (2014). Immigration: What about the children and grandchildren? In B. Chiswick & P. W. Miller (Eds.), *Handbook of the economics of international migration* (pp. 1142–1189). Elsevier.

Train, K. E. (2009). *Discrete choice methods with simulation*. Cambridge university press.

van den Berg, G. J., & Weynandt, M. A. (2012). Explaining differences between the expected and actual duration until return migration: Economic changes.

Van Hoorn, A., & Maseland, R. (2010). Cultural differences between east and west germany after 1991: Communist values versus economic performance? *Journal of Economic Behavior & Organization*, *76*(3), 791–804.

Watt, H. M. (2010). Gender and occupational choice. *Handbook of gender research in psychology* (pp. 379–400). Springer.

A. MORE ON DATA AND DESCRIPTIVE STATISTICS

Table (2) Summary statistics grouped by migration background.

	ISCO SKILL DATASET				LANGUAGE SKILL DATASET			
	NATIVE	DIRECT	INDIRECT	TOTAL	NATIVE	DIRECT	INDIRECT	TOTAL
Continuous variables								
Years of Education	12.97	10.78	12.46	12.62	12.97	10.78	12.46	12.62
	(2.720)	(1.971)	(2.924)	(2.746)	(2.721)	(1.972)	(2.922)	(2.747)
Work experience	16.10	13.52	12.60	15.53	15.97	13.47	12.55	15.41
	(7.968)	(7.894)	(7.077)	(7.994)	(7.944)	(7.877)	(7.155)	(7.972)
Language distance	0	87.70	0	12.99	0	87.71	0	12.99
	(0)	(12.06)	(0)	(31.50)	(0)	(12.04)	(0)	(31.50)
Years in Germany	0	15.66	11.73	2.951	0	15.60	11.72	2.941
	(0)	(8.371)	(17.08)	(7.817)	(0)	(8.355)	(17.06)	(7.795)
Categorical variables								
Elementary (1)	0.0324	0.154	0.0638	0.0521	0.0822	0.246	0.123	0.109
	(0.177)	(0.361)	(0.244)	(0.222)	(0.275)	(0.431)	(0.328)	(0.311)
Blue-collar (2)	0.149	0.278	0.183	0.170	0.0969	0.182	0.116	0.110
	(0.356)	(0.448)	(0.387)	(0.375)	(0.296)	(0.386)	(0.321)	(0.313)
White-collar (3)	0.127	0.169	0.172	0.136	0.198	0.224	0.238	0.204
	(0.333)	(0.375)	(0.377)	(0.342)	(0.398)	(0.417)	(0.426)	(0.403)
Technical (4)	0.306	0.211	0.271	0.290	0.309	0.213	0.279	0.293
	(0.461)	(0.408)	(0.445)	(0.454)	(0.462)	(0.409)	(0.449)	(0.455)
Professional (5)	0.386	0.187	0.310	0.352	0.314	0.136	0.243	0.284
	(0.487)	(0.390)	(0.463)	(0.478)	(0.464)	(0.343)	(0.429)	(0.451)
Female	0.519	0.499	0.476	0.514	0.519	0.500	0.475	0.514
	(0.500)	(0.500)	(0.500)	(0.500)	(0.500)	(0.500)	(0.500)	(0.500)
Education east	0.276	0.0101	0.0589	0.225	0.276	0.0104	0.0581	0.225
	(0.447)	(0.100)	(0.235)	(0.418)	(0.447)	(0.101)	(0.234)	(0.418)
Education abroad	0.0168	0.823	0.110	0.141	0.0170	0.822	0.110	0.141
	(0.129)	(0.382)	(0.312)	(0.348)	(0.129)	(0.382)	(0.313)	(0.348)
Married to German	0.556	0.0708	0.159	0.462	0.555	0.0708	0.160	0.462
	(0.497)	(0.257)	(0.366)	(0.499)	(0.497)	(0.256)	(0.366)	(0.499)
Children > 1	0.812	0.862	0.763	0.817	0.812	0.862	0.763	0.817
	(0.391)	(0.345)	(0.425)	(0.387)	(0.390)	(0.345)	(0.425)	(0.387)
East	0.267	0.0765	0.0744	0.228	0.268	0.0770	0.0713	0.229
	(0.442)	(0.266)	(0.263)	(0.420)	(0.443)	(0.267)	(0.257)	(0.420)
Germanic	1	0.0277	0.198	0.813	1	0.0274	0.198	0.813
	(0)	(0.164)	(0.399)	(0.390)	(0)	(0.163)	(0.399)	(0.390)
Western European	0	0.0907	0.141	0.0211	0	0.0913	0.143	0.0212
	(0)	(0.287)	(0.349)	(0.144)	(0)	(0.288)	(0.350)	(0.144)
Eastern European	0	0.726	0.297	0.124	0	0.726	0.297	0.123
	(0)	(0.446)	(0.457)	(0.329)	(0)	(0.446)	(0.457)	(0.329)
Turkish/Greek	0	0.156	0.364	0.0427	0	0.156	0.362	0.0425
	(0)	(0.363)	(0.481)	(0.202)	(0)	(0.363)	(0.481)	(0.202)
Baby boom	0.383	0.141	0.0957	0.332	0.384	0.141	0.0966	0.333
	(0.486)	(0.348)	(0.294)	(0.471)	(0.486)	(0.348)	(0.296)	(0.471)
Gen. X	0.359	0.330	0.365	0.355	0.359	0.330	0.366	0.355
	(0.480)	(0.470)	(0.482)	(0.478)	(0.480)	(0.470)	(0.482)	(0.478)
Millennial	0.258	0.529	0.540	0.313	0.257	0.529	0.537	0.313
	(0.437)	(0.499)	(0.499)	(0.464)	(0.437)	(0.499)	(0.499)	(0.464)
Elementary (1) [Parent]	0.0309	0.113	0.147	0.0493	0.161	0.268	0.355	0.187
	(0.173)	(0.317)	(0.354)	(0.217)	(0.367)	(0.443)	(0.479)	(0.390)
Blue-collar (2) [Parent]	0.309	0.369	0.397	0.323	0.159	0.212	0.155	0.167
	(0.462)	(0.483)	(0.489)	(0.468)	(0.366)	(0.409)	(0.362)	(0.373)
White-collar (3) [Parent]	0.130	0.135	0.0932	0.129	0.215	0.210	0.231	0.215
	(0.336)	(0.342)	(0.291)	(0.335)	(0.410)	(0.408)	(0.422)	(0.411)
Technical (4) [Parent]	0.277	0.192	0.239	0.262	0.268	0.158	0.163	0.246
	(0.447)	(0.394)	(0.426)	(0.440)	(0.443)	(0.365)	(0.369)	(0.431)
Professional (5) [Parent]	0.253	0.191	0.124	0.237	0.198	0.151	0.0966	0.186
	(0.435)	(0.393)	(0.330)	(0.425)	(0.398)	(0.358)	(0.296)	(0.389)
N. of observations	18108	3361	1223	**22692**	18120	3363	1221	**22704**

Reported are Grouped Means (Continuous Variables) and Grouped Proportions (Categorical Variables).
Standard Errors in parentheses.

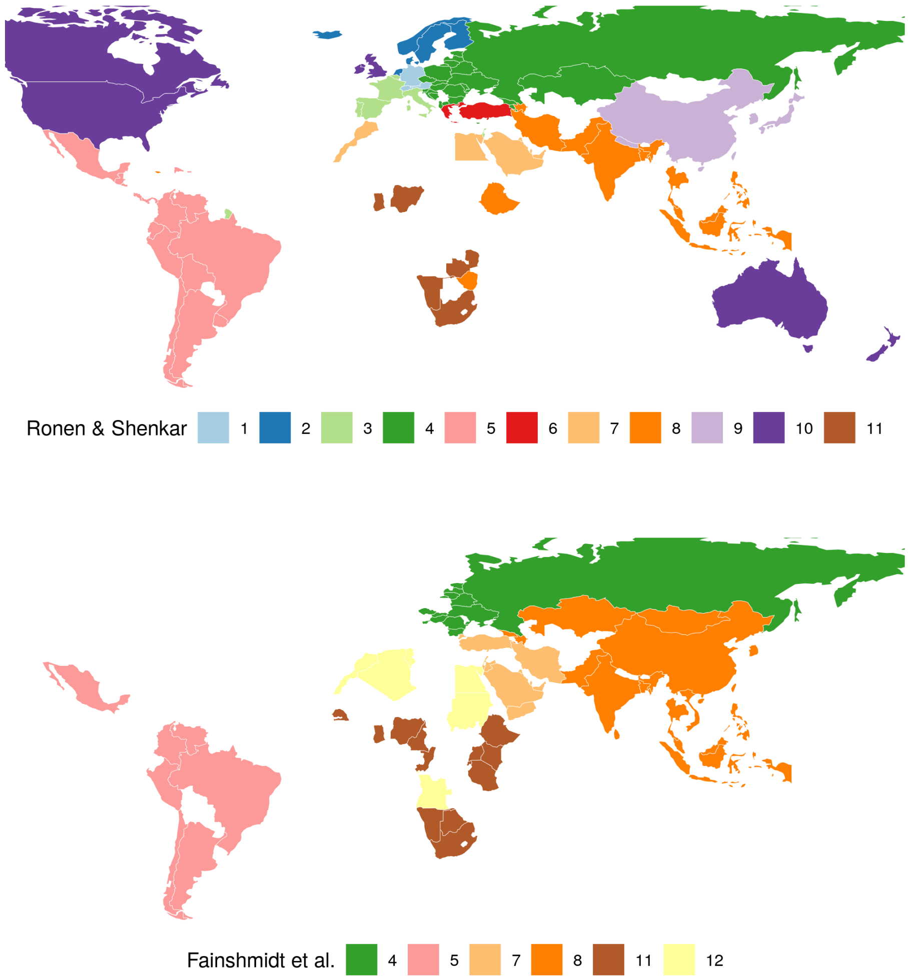


Figure (8) Cultural Maps. (1) Germanic (2) Scandinavian (3) Western European (4) Eastern European (5) Latin American (6) Turkish/Greek (7) Arabic (8) Asian (9) Confucian (10) Anglo-Saxon (11) Sub-saharan Africa (12) North African. (Top) Ronen and Shenkar (2013) (Bottom) Fainshmidt et al. (2018). Note: Angola seems to be misclassified by Fainshmidt et al. (2018).

Figure (9) Labor Force Participation by migration background.

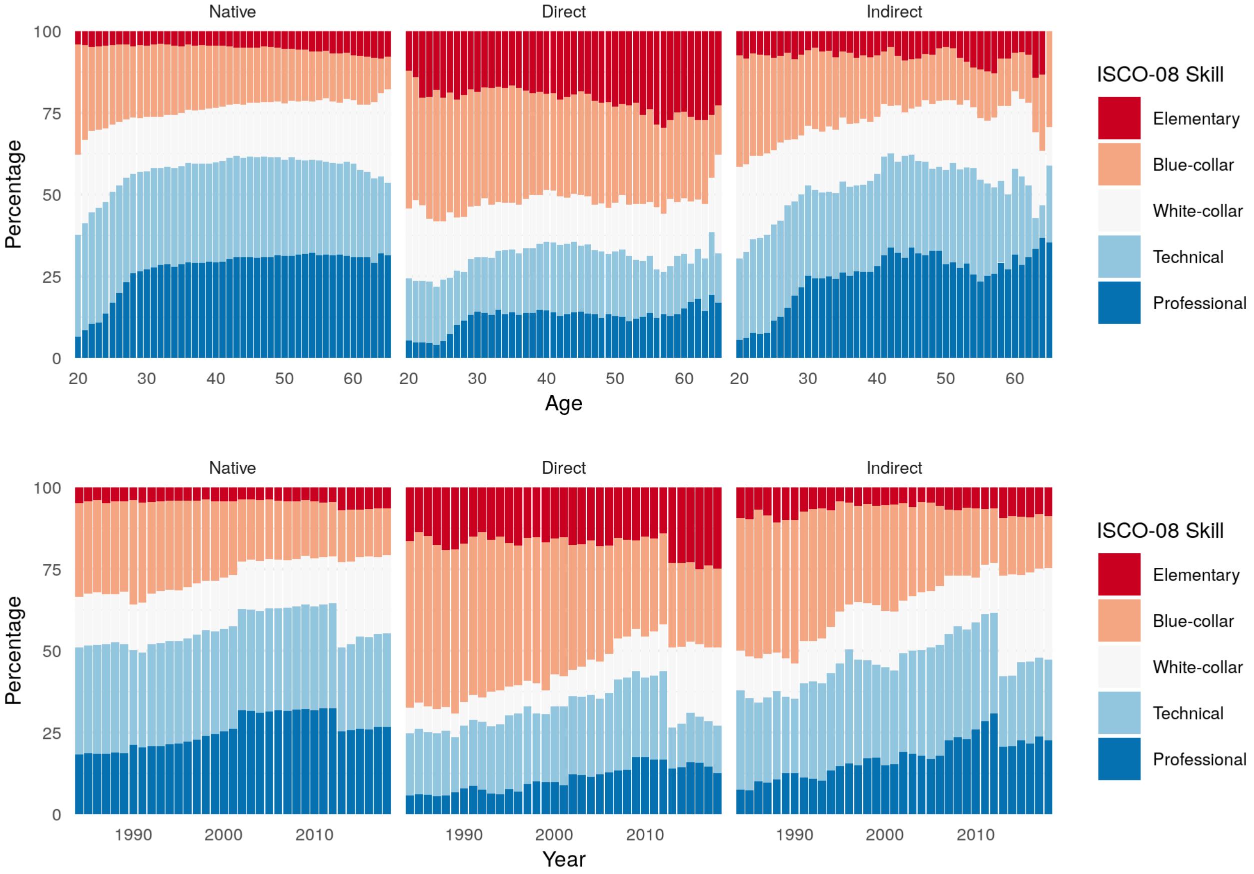


Figure (10) ISCO Skill Distribution by migration background.

Figure (11) Language Skill Distribution by migration background. Note: *Skill 1* is the language skill using only "Knowledge of the language". *Skill 2* is the weighted average of Reading, Listening, Writing, Speaking and Knowledge of the language.

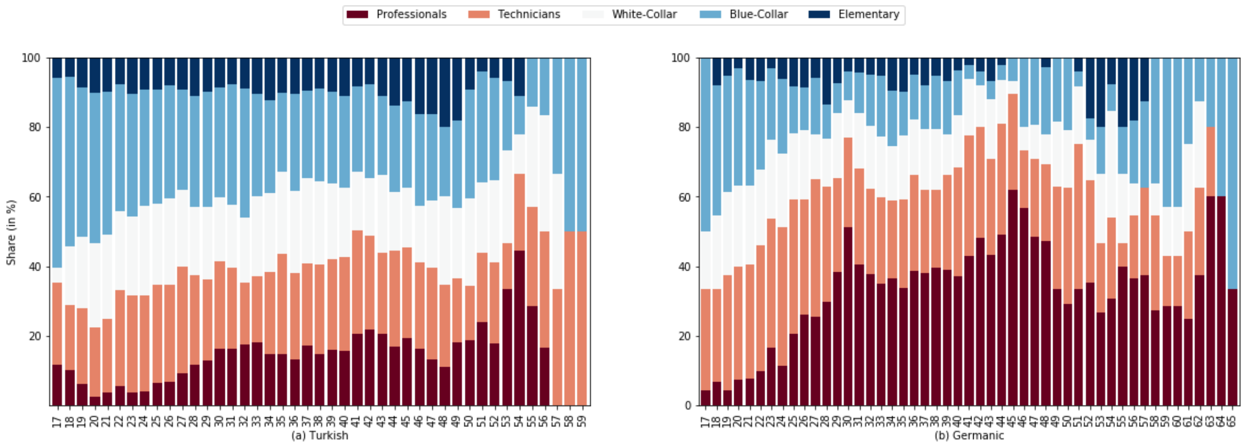


Figure (12) ISCO Skill Distribution by Turkish and Germanic origin.

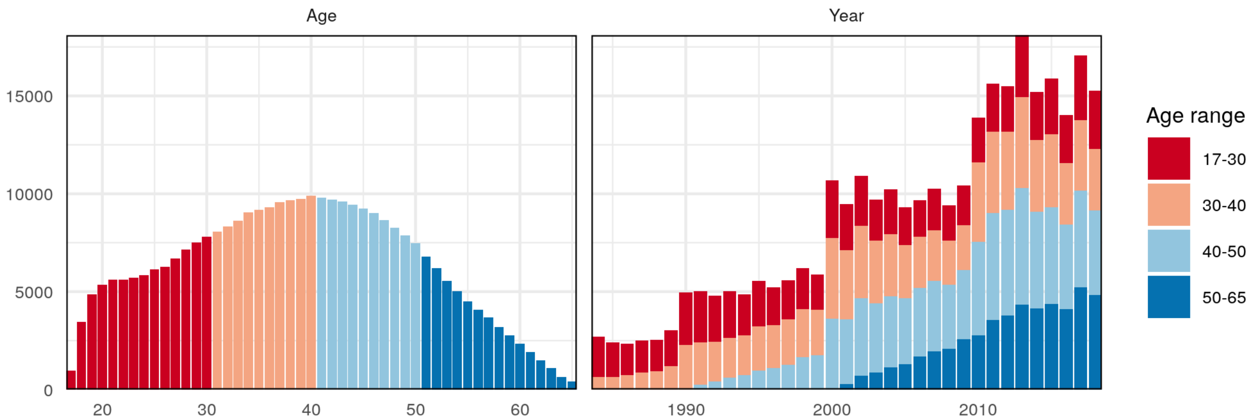


Figure (13) Number of observations by age and year for the complete sample without filtering. Note: We took only observations from 30 to 50 years old, across 1984-2018.

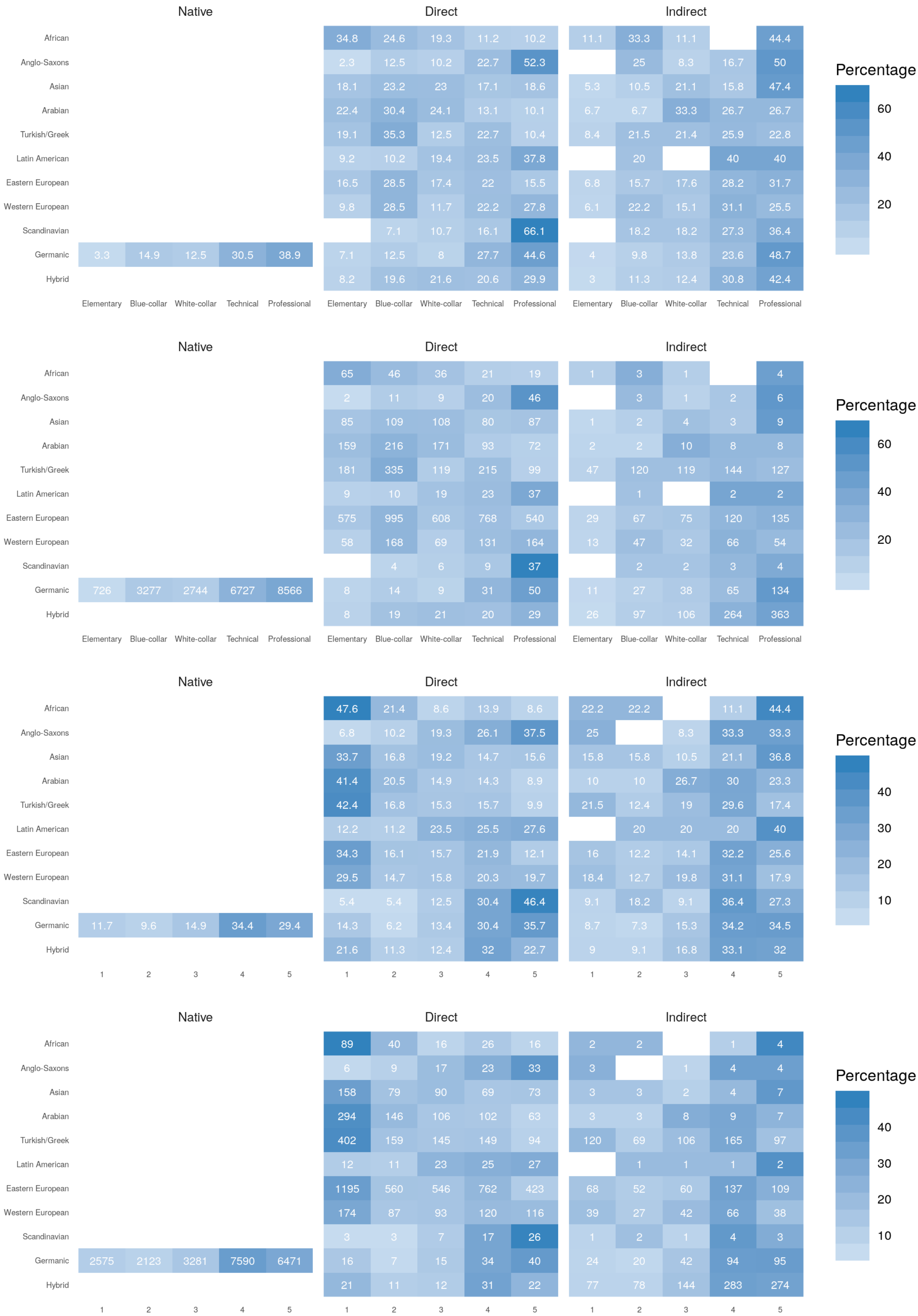


Figure (14) Occupational decisions by culture (Top 2) ISCO skill (Bottom 2) Language-interactivity skill.

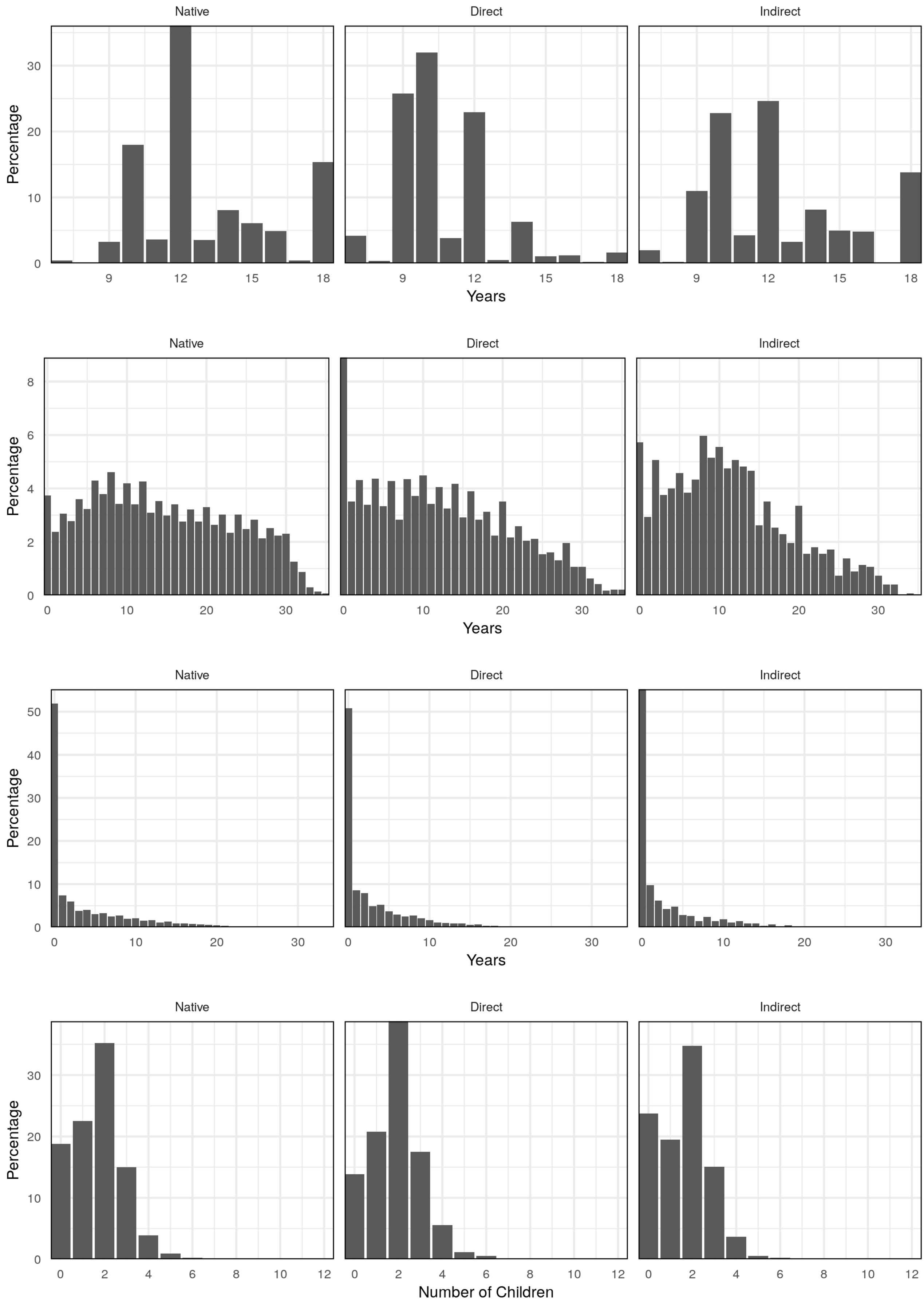


Figure (15) Distribution differences across migration backgrounds (1) Years of education (2) Full-time experience (3) Part-time experience. (4) Number of Children [*Fertility*]. For all plots, the reported time was rounded to the nearest integer.

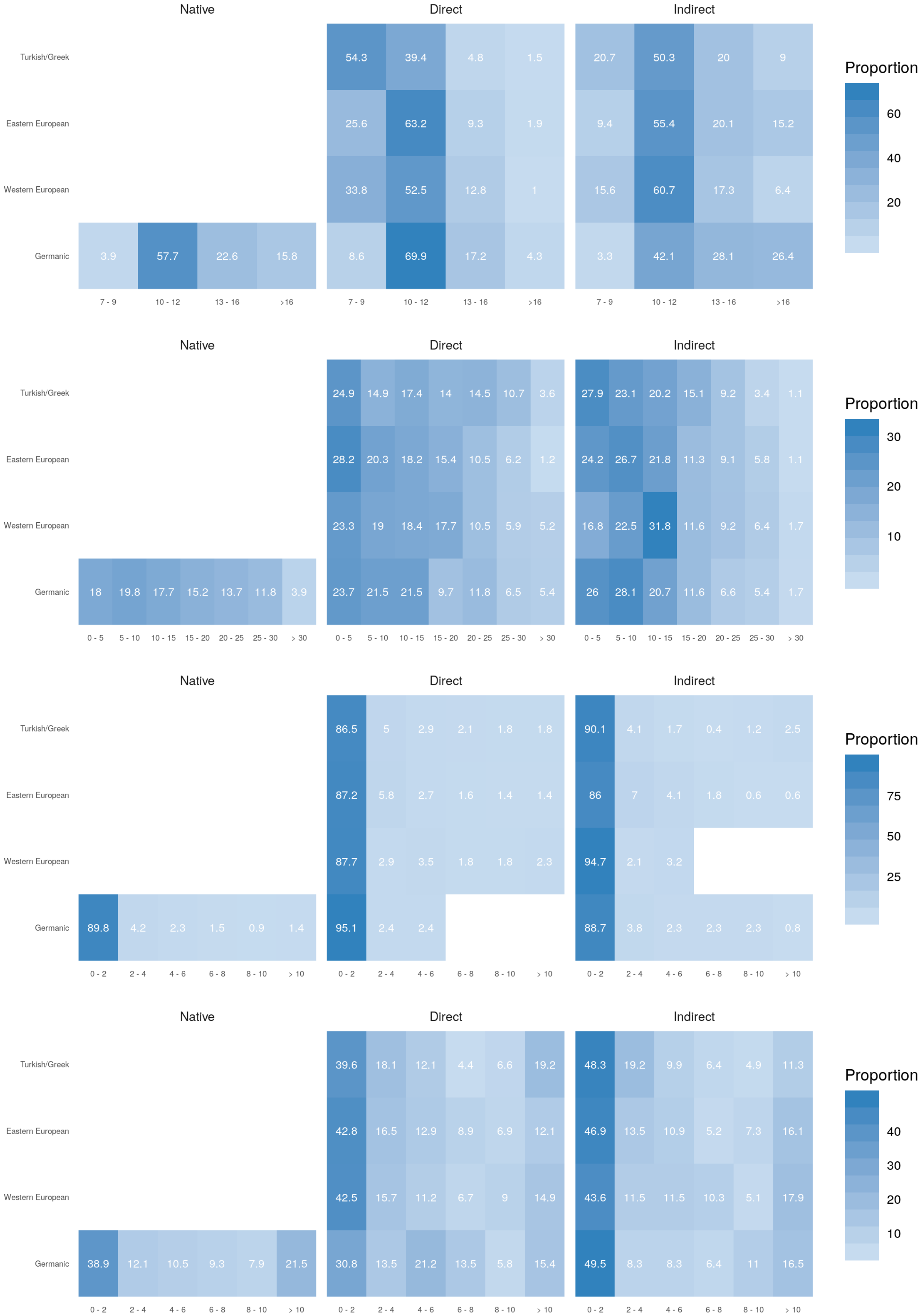


Figure (16) Distribution differences across cultures and migration backgrounds (1) Years of education (2) Full-time experience (3) Part-time experience for males (4) Part-time experience for females.

Figure (17) ISCO Skill Distribution differences across cohorts (Top) Dis-aggregated by culture and migration background (Bottom) Dis-aggregated by migration background.

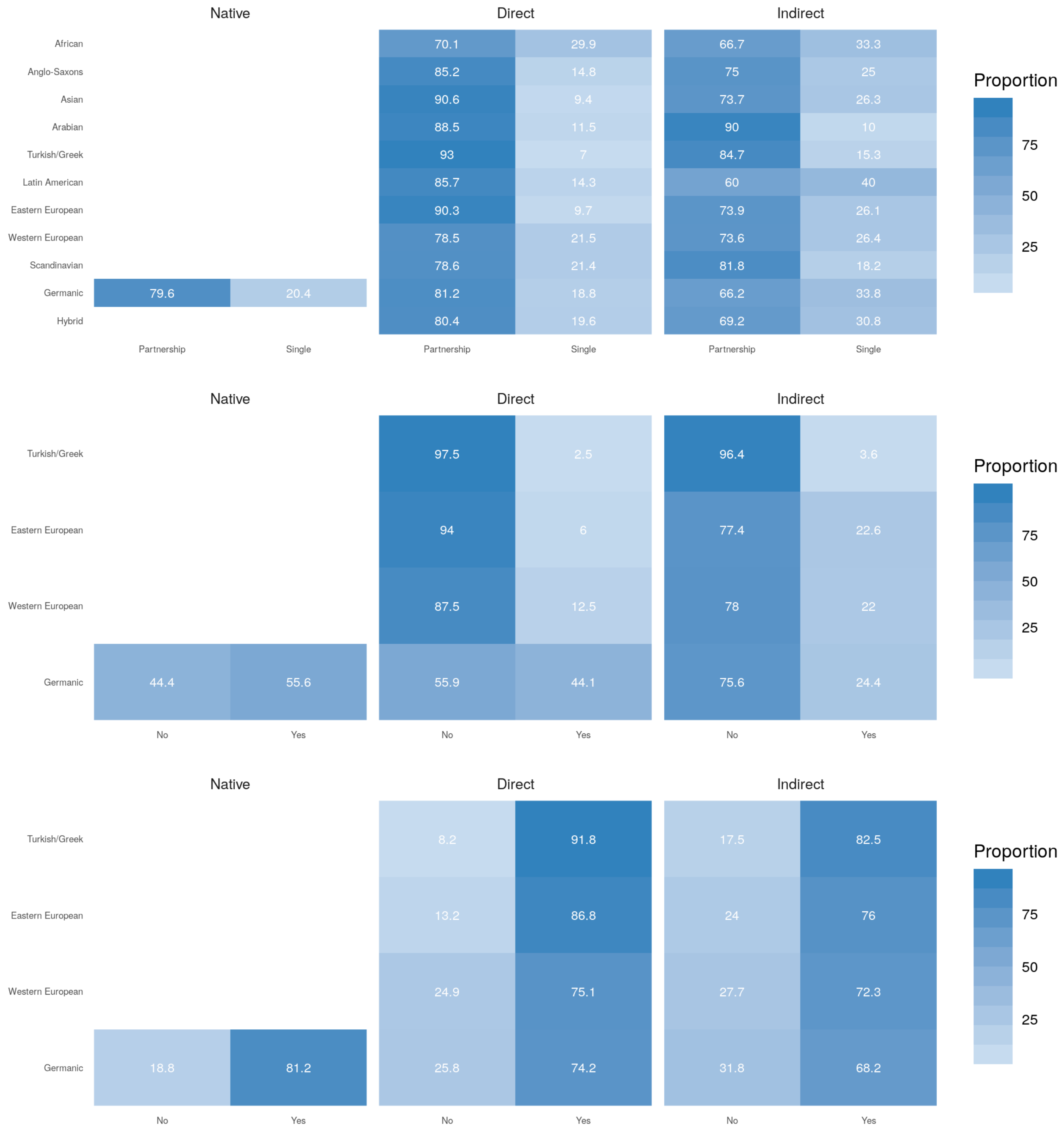


Figure (18) Distribution differences across cultures (Top) Single or with a registered Partner at least once during life-time (Middle) Married to German. (Bottom) Children > 0

Figure (19) ISCO Skill Distribution differences across migration backgrounds (Top) Parental Background (Bottom) Parental Background by culture. Note: Parental occupation is on y axis, individual occupation is on x axis.

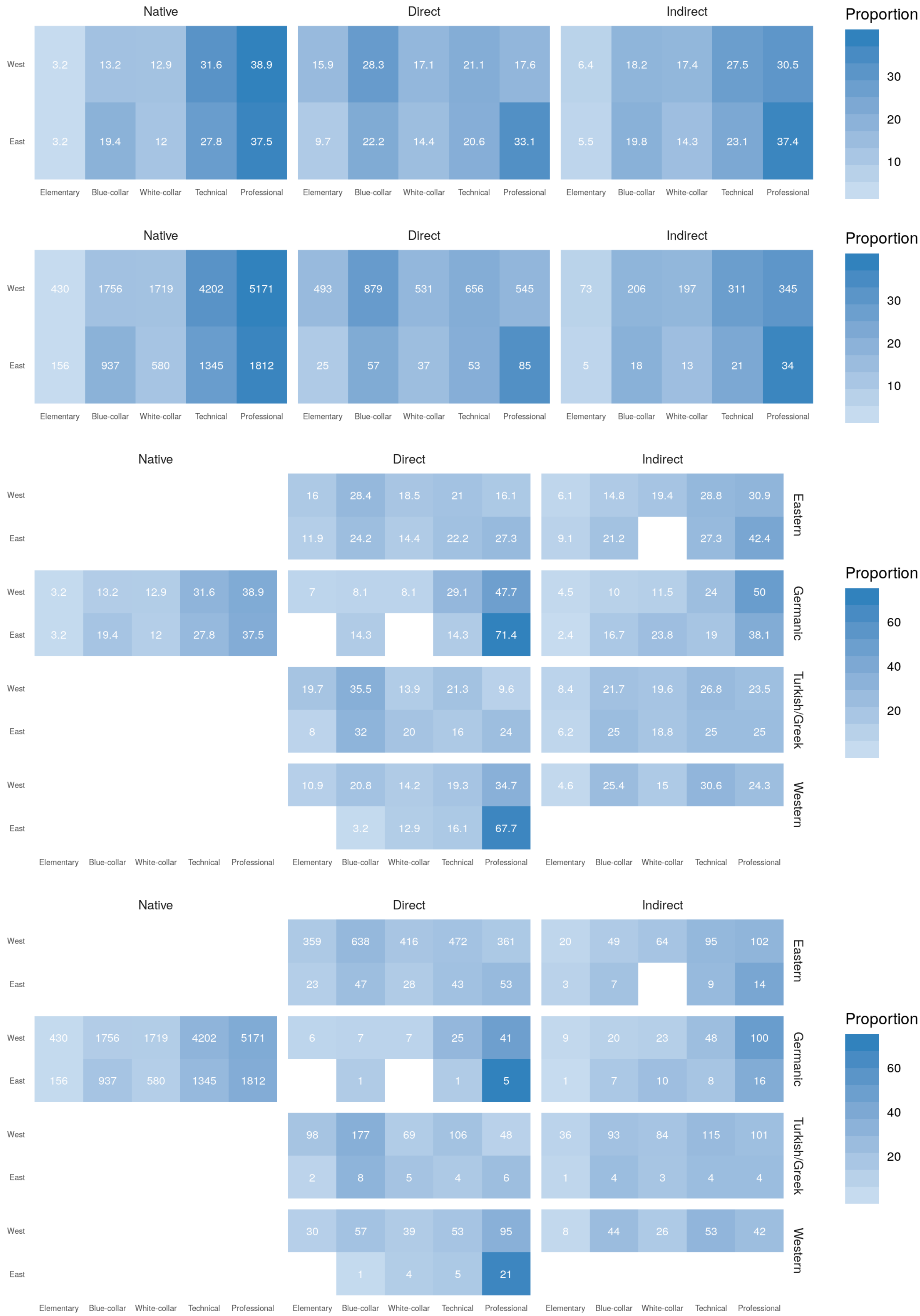


Figure (20) Distribution differences across migration backgrounds (1) Location within Germany (2) Location within Germany [N. of observations] (3) Location within Germany by cultures (4) Location within Germany by cultures [N. of observations]

B. COMPLEMENTARY RESULTS

Table (3) Average Marginal Effects estimated by the Basic Model for ISCO Skill

	ELEMENTARY	BLUE-COLLAR	WHITE-COLLAR	TECHNICAL	PROFESSIONAL
Female	0.0257***	-0.228***	0.0812***	0.158***	-0.0370***
	(0.00306)	(0.00480)	(0.00474)	(0.00612)	(0.00543)
Years of education	-0.0250***	-0.0471***	-0.0153***	-0.000143	0.0875***
	(0.00113)	(0.00144)	(0.00115)	(0.00123)	(0.000737)
Work experience	-0.00285***	0.00166***	-0.00334***	0.00124**	0.00329***
	(0.000251)	(0.000285)	(0.000340)	(0.000398)	(0.000355)
Direct migrant	0.0276*	-0.0383	-0.0285	-0.0863***	0.125***
	(0.0138)	(0.0202)	(0.0180)	(0.0240)	(0.0256)
Indirect migrant	0.00814	-0.0370	0.00779	-0.0167	0.0378
	(0.0113)	(0.0205)	(0.0209)	(0.0253)	(0.0215)
Western European	-0.0133	0.0109	0.00251	-0.0238	0.0237
	(0.0110)	(0.0256)	(0.0241)	(0.0319)	(0.0297)
Eastern European	0.00901	0.0813**	0.0364	-0.0200	-0.107***
	(0.0129)	(0.0265)	(0.0236)	(0.0286)	(0.0214)
Turkish/Greek	0.00392	0.0444	0.0307	0.0256	-0.105***
	(0.0126)	(0.0260)	(0.0246)	(0.0318)	(0.0232)

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Delta Method Standard Errors in parentheses.

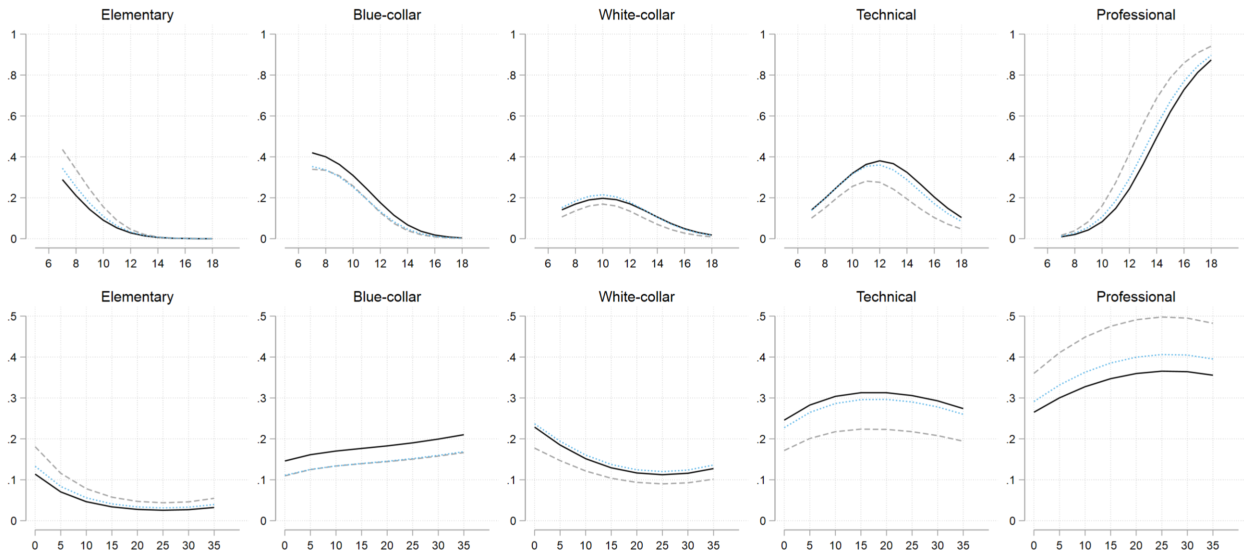


Figure (21) Average Marginal Effects of the continuous variables estimated in the basic model for ISCO Skill discriminated by migrant status. (Top) Years of education. (Bottom) Years of experience. In addition, Solid = Native, Dotted = Indirect Migrant, Dashed = Direct Migrant

Table (4) Average Marginal Effects for Basic model's variation for ISCO Skill

	(1)					(2)					(3)				
	ELEM.	B-C.	W-C.	TECH.	PRO.	ELEM.	B-C.	W-C.	TECH.	PRO.	ELEM.	B-C.	W-C.	TECH.	PRO.
Female	0.0257***	-0.228***	0.0812***	0.158***	-0.0370***	0.0257***	-0.228***	0.0812***	0.158***	-0.0374***	0.0257***	-0.228***	0.0812***	0.158***	-0.0371***
	(0.00306)	(0.00480)	(0.00474)	(0.00612)	(0.00543)	(0.00307)	(0.00480)	(0.00474)	(0.00612)	(0.00543)	(0.00306)	(0.00480)	(0.00474)	(0.00612)	(0.00543)
Years of education	-0.0250***	-0.0471***	-0.0153***	-0.000143	0.0875***	-0.0250***	-0.0470***	-0.0152***	-0.000189	0.0873***	-0.0250***	-0.0469***	-0.0152***	-0.000207	0.0873***
	(0.00113)	(0.00144)	(0.00115)	(0.00123)	(0.000737)	(0.00112)	(0.00144)	(0.00115)	(0.00123)	(0.000737)	(0.00113)	(0.00144)	(0.00115)	(0.00123)	(0.000736)
Work experience	-0.00285***	0.00166***	-0.00334***	0.00124**	0.00329***	-0.00285***	0.00168***	-0.00333***	0.00128**	0.00323***	-0.00285***	0.00166***	-0.00333***	0.00124**	0.00329***
	(0.000251)	(0.000285)	(0.000340)	(0.000398)	(0.000355)	(0.000251)	(0.000285)	(0.000340)	(0.000397)	(0.000356)	(0.000251)	(0.000285)	(0.000340)	(0.000398)	(0.000355)
Western European	-0.0133	0.0109	0.00251	-0.0238	0.0237	-0.00730	-0.0316*	0.00309	-0.0442*	0.0800***	-0.0185	-0.0187	-0.0116	-0.0249	0.0737*
	(0.0110)	(0.0256)	(0.0241)	(0.0319)	(0.0297)	(0.00743)	(0.0140)	(0.0165)	(0.0220)	(0.0230)	(0.0111)	(0.0245)	(0.0232)	(0.0327)	(0.0328)
Eastern European	0.00901	0.0813**	0.0364	-0.0200	-0.107***	0.0174*	0.0313*	0.0407**	-0.0318	-0.0576***	0.00255	0.0447	0.0218	-0.0113	-0.0578*
	(0.0129)	(0.0265)	(0.0236)	(0.0286)	(0.0214)	(0.00777)	(0.0124)	(0.0134)	(0.0174)	(0.0157)	(0.0135)	(0.0269)	(0.0239)	(0.0306)	(0.0251)
Turkish/Greek	0.00392	0.0444	0.0307	0.0256	-0.105***	0.0109	0.000295	0.0387**	0.0142	-0.0641***	-0.00262	0.00291	0.0149	0.0395	-0.0547*
	(0.0126)	(0.0260)	(0.0246)	(0.0318)	(0.0232)	(0.00766)	(0.0121)	(0.0142)	(0.0189)	(0.0169)	(0.0135)	(0.0258)	(0.0248)	(0.0340)	(0.0275)
Direct migrant	0.0276*	-0.0383	-0.0285	-0.0863***	0.125***						0.00984	-0.140***	-0.0758*	-0.0740	0.280***
	(0.0138)	(0.0202)	(0.0180)	(0.0240)	(0.0256)						(0.0206)	(0.0288)	(0.0306)	(0.0460)	(0.0449)
Indirect migrant	0.00814	-0.0370	0.00779	-0.0167	0.0378						0.0134	-0.00977	0.0206	-0.0223	-0.00189
	(0.0113)	(0.0205)	(0.0209)	(0.0253)	(0.0215)						(0.0132)	(0.0235)	(0.0231)	(0.0250)	(0.0212)
Language distance						0.000219**	0.000128	-0.000299*	-0.000772***	0.000725***	0.000170	0.00213**	0.000865	-0.000805	-0.00236***
						(0.0000710)	(0.000125)	(0.000128)	(0.000198)	(0.000187)	(0.000311)	(0.000744)	(0.000661)	(0.000797)	(0.000579)

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Clustered Standard Errors at the Household level in parentheses.

Table (5) Average Marginal Effects: Second generation - Native for ISCO Skill Level.

	ELEMENTARY	BLUE-COLLAR	WHITE-COLLAR	TECHNICAL	PROFESSIONAL
Female	0.0197***	-0.249***	0.120***	0.153***	-0.0436***
	(-0.00318)	(-0.00466)	(-0.00522)	(-0.00598)	(-0.00492)
Years of education	-0.0306***	-0.0445***	-0.0141***	0.0151***	0.0741***
	(-0.00134)	(-0.00131)	(-0.00108)	(-0.00106)	(-0.000559)
Work experience	-0.00300***	0.000731**	-0.00234***	0.00197***	0.00264***
	(-0.00024)	(-0.000277)	(-0.000341)	(-0.000382)	(-0.000323)
Germanic	-0.0045	-0.0153	-0.00348	-0.0238	0.0471*
	(-0.0137)	(-0.0223)	(-0.0244)	(-0.0273)	(-0.021)
Western European	-0.0079	0.0214	0.00398	0.00356	-0.0211
	(-0.0117)	(-0.0194)	(-0.0244)	(-0.0322)	(-0.0301)
Eastern European	0.0104	-0.0188	0.0272	-0.0265	0.00773
	(-0.0107)	(-0.0149)	(-0.0196)	(-0.0226)	(-0.019)
Turkish/Greek	0.00703	0.00304	0.0705***	-0.0443*	-0.0362
	(0.00809)	(0.0138)	(-0.0179)	(-0.0195)	(-0.0185)

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Delta Method Standard Errors in parentheses.

Table (6) Average Marginal Effects calculated by the Augmented Model for ISCO Skill

	ELEMENTARY	BLUE-COLLAR	WHITE-COLLAR	TECHNICAL	PROFESSIONAL
Female	0.0240***	-0.224***	0.0860***	0.155***	-0.0412***
	(0.00305)	(0.00493)	(0.00479)	(0.00623)	(0.00550)
Years of education	-0.0229***	-0.0447***	-0.0143***	-0.000346	0.0822***
	(0.00119)	(0.00158)	(0.00124)	(0.00135)	(0.000852)
Work experience	-0.00309***	0.00251***	-0.00201***	0.000769	0.00182***
	(0.000261)	(0.000342)	(0.000385)	(0.000466)	(0.000417)
Direct migrant	0.0511*	-0.0299	-0.0171	-0.124***	0.120***
	(0.0206)	(0.0249)	(0.0240)	(0.0280)	(0.0329)
Indirect migrant	0.0135	-0.0286	0.00161	-0.0412	0.0547*
	(0.0118)	(0.0213)	(0.0202)	(0.0260)	(0.0226)
Western European	-0.0155	0.0103	0.00639	-0.0317	0.0306
	(0.0110)	(0.0257)	(0.0244)	(0.0316)	(0.0288)
Eastern European	0.00818	0.0742**	0.0314	-0.0322	-0.0815***
	(0.0134)	(0.0264)	(0.0233)	(0.0280)	(0.0214)
Turkish/Greek	-0.00137	0.0353	0.0272	0.0129	-0.0740**
	(0.0124)	(0.0257)	(0.0244)	(0.0314)	(0.0238)
Education east	0.0176*	0.0613***	0.0187	-0.0333**	-0.0642***
	(0.00715)	(0.00985)	(0.00961)	(0.0116)	(0.00962)
Education abroad	0.0218**	0.000997	-0.0281**	-0.0379*	0.0431**
	(0.00716)	(0.0112)	(0.0105)	(0.0155)	(0.0149)
Years in Germany	-0.00291***	0.000246	0.000171	0.00798***	-0.00549**
	(0.000610)	(0.00128)	(0.00138)	(0.00209)	(0.00204)
Married to German	-0.0100**	-0.0143**	-0.0134**	0.0110	0.0268***
	(0.00331)	(0.00516)	(0.00498)	(0.00640)	(0.00574)
Children > 0	0.00654	0.00607	-0.00444	-0.00928	0.00111
	(0.00388)	(0.00585)	(0.00619)	(0.00788)	(0.00692)
Blue-collar parent	-0.0207***	0.0200*	-0.0292**	0.0170	0.0128
	(0.00622)	(0.00958)	(0.0110)	(0.0146)	(0.0137)
White-collar parent	-0.0197**	-0.0175	-0.00236	0.0198	0.0198
	(0.00695)	(0.0106)	(0.0123)	(0.0158)	(0.0150)
Technician parent	-0.0367***	-0.0233*	-0.0355**	0.0474**	0.0481***
	(0.00645)	(0.00995)	(0.0113)	(0.0149)	(0.0141)
Professional parent	-0.0409***	-0.0462***	-0.0398***	0.0290	0.0979***
	(0.00692)	(0.0106)	(0.0119)	(0.0155)	(0.0146)
Gen. X	-0.000170	0.0209***	0.0230***	0.00344	-0.0471***
	(0.00395)	(0.00528)	(0.00547)	(0.00729)	(0.00650)
Millennial	-0.00807	0.0360***	0.0489***	-0.0135	-0.0633***
	(0.00430)	(0.00696)	(0.00676)	(0.00881)	(0.00779)
East = 1	-0.00791	0.0148	-0.0276**	-0.0212	0.0418***
	(0.00529)	(0.00917)	(0.00841)	(0.0117)	(0.0104)

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Delta Method Standard Errors in parentheses.

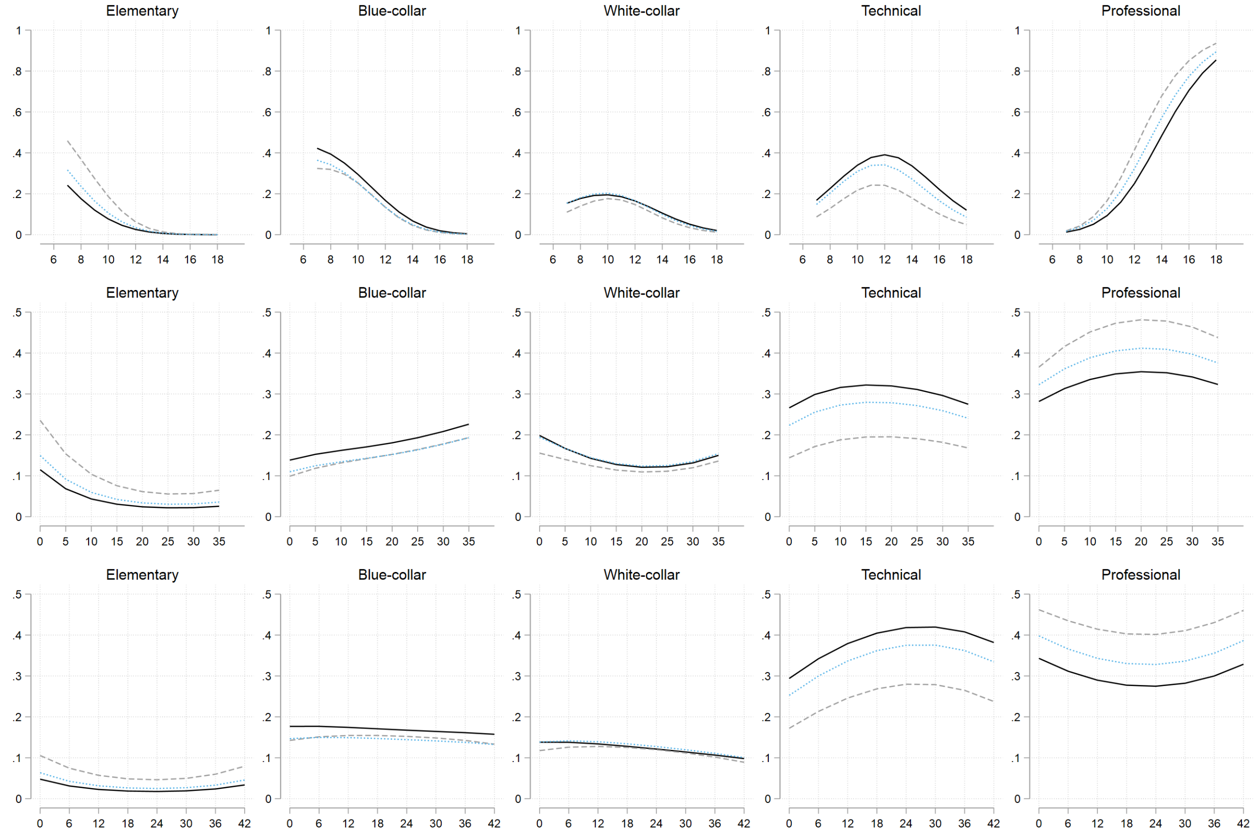


Figure (22) Average Marginal Effects of the continuous variables estimated in the augmented model for ISCO Skill discriminated by migrant status. (Top) Years of education. (Middle) Years of experience (Bottom) Years in Germany. In addition, Solid = Native, Dotted = Indirect Migrant, Dashed = Direct Migrant

Table (7) Average Marginal Effects estimated by the Basic Model for Language Skill

	1	2	3	4	5
Female	-0.0572***	-0.0746***	0.00985	0.0836***	0.0384***
	(-0.00419)	(-0.00435)	(-0.00554)	(-0.00628)	(-0.00548)
Years of education	-0.0463***	-0.0217***	-0.0181***	0.0195***	0.0667***
	(-0.0015)	(-0.0012)	(-0.00121)	(-0.00113)	(-0.000776)
Work experience	-0.00118***	-0.000898**	-0.00129***	0.000898*	0.00246***
	(-0.000292)	(-0.000289)	(-0.000362)	(-0.00041)	(-0.000368)
Direct migrant	0.0289	-0.032	-0.0296	0.0352	-0.00241
	(-0.0207)	(-0.0201)	(-0.0236)	(-0.0284)	(-0.0274)
Indirect migrant	0.00692	-0.0422*	0.000543	0.0618*	-0.027
	(-0.0185)	(-0.0186)	(-0.024)	(-0.0262)	(-0.0222)
Western European	-0.0305	0.0378	0.0191	-0.0761*	0.0497
	(-0.0172)	(-0.0283)	(-0.0306)	(-0.0296)	(-0.0332)
Eastern European	0.0208	0.0776*	0.0143	-0.0872***	-0.0255
	(-0.0213)	(-0.0303)	(-0.0271)	(-0.0246)	(-0.0257)
Turkish/Greek	-0.00501	0.0587*	0.0452	-0.0821**	-0.0169
	(-0.0191)	(-0.0289)	(-0.0297)	(-0.0265)	(-0.028)

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Delta Method Standard Errors in parentheses.

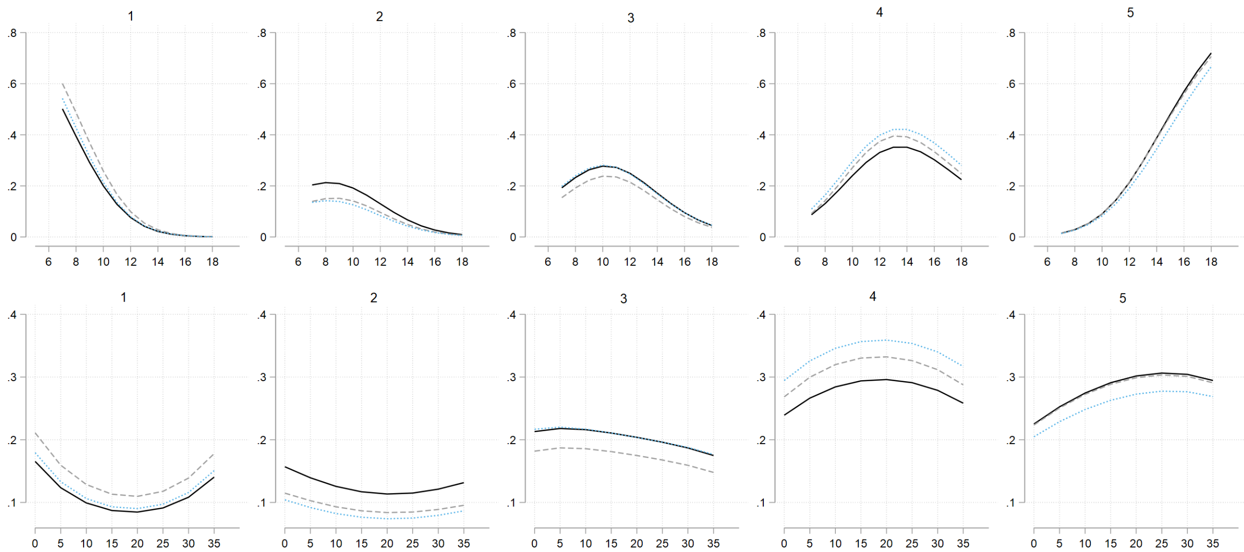


Figure (23) Average Marginal Effects of the continuous variables estimated in the basic model for Language Skill discriminated by migrant status. (Top) Years of education. (Bottom) Years of experience. In addition, Solid = Native, Dotted = Indirect Migrant, Dashed = Direct Migrant

Table (8) Average Marginal Effects for Basic model's variation for Language Skill Level

	(1)					(2)					(3)				
	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
Female	-0.0572***	-0.0746***	0.00985	0.0836***	0.0384***	-0.0572***	-0.0744***	0.00990	0.0832***	0.0386***	-0.0572***	-0.0745***	0.00991	0.0836***	0.0382***
	(0.00419)	(0.00435)	(0.00554)	(0.00628)	(0.00548)	(0.00419)	(0.00435)	(0.00554)	(0.00628)	(0.00548)	(0.00419)	(0.00435)	(0.00554)	(0.00628)	(0.00547)
Years of education	-0.0463***	-0.0217***	-0.0181***	0.0195***	0.0667***	-0.0462***	-0.0217***	-0.0181***	0.0195***	0.0665***	-0.0462***	-0.0217***	-0.0181***	0.0194***	0.0666***
	(0.00150)	(0.00120)	(0.00121)	(0.00113)	(0.000776)	(0.00150)	(0.00120)	(0.00121)	(0.00114)	(0.000776)	(0.00150)	(0.00120)	(0.00121)	(0.00113)	(0.000776)
Work experience	-0.00118***	-0.000898**	-0.00129***	0.000898*	0.00246***	-0.00118***	-0.000875**	-0.00128***	0.000847*	0.00249***	-0.00117***	-0.000896**	-0.00129***	0.000895*	0.00246***
	(0.000292)	(0.000289)	(0.000362)	(0.000410)	(0.000368)	(0.000292)	(0.000289)	(0.000362)	(0.000410)	(0.000368)	(0.000292)	(0.000289)	(0.000362)	(0.000410)	(0.000368)
Western European	-0.0305	0.0378	0.0191	-0.0761*	0.0497	-0.0270*	-0.00873	0.0189	-0.0227	0.0395	-0.0447**	0.0198	0.00538	-0.0724*	0.0919**
	(0.0172)	(0.0283)	(0.0306)	(0.0296)	(0.0332)	(0.0110)	(0.0132)	(0.0202)	(0.0242)	(0.0241)	(0.0168)	(0.0262)	(0.0303)	(0.0312)	(0.0353)
Eastern European	0.0208	0.0776*	0.0143	-0.0872***	-0.0255	0.0256*	0.0203	0.0159	-0.0328	-0.0290	0.0000384	0.0554	0.00319	-0.0791**	0.0205
	(0.0213)	(0.0303)	(0.0271)	(0.0246)	(0.0257)	(0.0112)	(0.0116)	(0.0152)	(0.0177)	(0.0163)	(0.0212)	(0.0292)	(0.0277)	(0.0273)	(0.0285)
Turkish/Greek	-0.00501	0.0587*	0.0452	-0.0821**	-0.0169	-0.00131	0.00631	0.0501**	-0.0256	-0.0296	-0.0263	0.0347	0.0327	-0.0714*	0.0303
	(0.0191)	(0.0289)	(0.0297)	(0.0265)	(0.0280)	(0.0104)	(0.0111)	(0.0163)	(0.0193)	(0.0186)	(0.0191)	(0.0277)	(0.0308)	(0.0295)	(0.0316)
Direct migrant	0.0289	-0.0320	-0.0296	0.0352	-0.00241						-0.0418	-0.0768*	-0.0695	0.0321	0.156**
	(0.0207)	(0.0201)	(0.0236)	(0.0284)	(0.0274)						(0.0277)	(0.0357)	(0.0403)	(0.0530)	(0.0549)
Indirect migrant	0.00692	-0.0422*	0.000543	0.0618*	-0.0270						0.0300	-0.0338	0.00875	0.0501	-0.0550**
	(0.0185)	(0.0186)	(0.0240)	(0.0262)	(0.0222)						(0.0236)	(0.0213)	(0.0255)	(0.0275)	(0.0207)
Language distance						0.000290**	0.000175	-0.000290	-0.000220	0.0000449	0.00109*	0.000912	0.000610	-0.000354	-0.00226***
						(0.000107)	(0.000113)	(0.000158)	(0.000207)	(0.000197)	(0.000468)	(0.000652)	(0.000682)	(0.000712)	(0.000619)

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Clustered Standard Errors at the Household level in parentheses.

Table (9) Average Marginal Effects for the Augmented Model for Language Skill Level

	1	2	3	4	5
Female	-0.0610***	-0.0696***	0.0125*	0.0801***	0.0380***
	(0.00425)	(0.00444)	(0.00564)	(0.00637)	(0.00557)
Years of education	-0.0443***	-0.0199***	-0.0163***	0.0171***	0.0634***
	(0.00160)	(0.00129)	(0.00129)	(0.00123)	(0.000870)
Work experience	-0.00155***	0.0000643	-0.000476	-0.0000599	0.00202***
	(0.000325)	(0.000338)	(0.000421)	(0.000477)	(0.000432)
Direct migrant	0.0530*	-0.0314	-0.0443	0.0228	-0.0000461
	(0.0268)	(0.0233)	(0.0286)	(0.0368)	(0.0344)
Indirect migrant	0.0263	-0.0393*	-0.0149	0.0528	-0.0250
	(0.0201)	(0.0190)	(0.0243)	(0.0273)	(0.0232)
Western European	-0.0365*	0.0399	0.0147	-0.0749*	0.0568
	(0.0167)	(0.0284)	(0.0304)	(0.0296)	(0.0331)
Eastern European	0.0161	0.0704*	0.000344	-0.0813**	-0.00551
	(0.0211)	(0.0294)	(0.0263)	(0.0248)	(0.0262)
Turkish/Greek	-0.0135	0.0536	0.0294	-0.0750**	0.00550
	(0.0184)	(0.0281)	(0.0290)	(0.0271)	(0.0290)
Education east	0.0622***	0.0197*	0.0118	-0.0399***	-0.0537***
	(0.0103)	(0.00911)	(0.0112)	(0.0120)	(0.00978)
Education abroad	0.0344**	-0.00173	-0.0325*	-0.0261	0.0259
	(0.0106)	(0.00937)	(0.0138)	(0.0159)	(0.0147)
Years in Germany	-0.00287**	0.000519	0.00472**	0.00228	-0.00465*
	(0.000989)	(0.00119)	(0.00168)	(0.00234)	(0.00224)
Married to German	-0.0143**	-0.00709	0.00272	0.00239	0.0163**
	(0.00463)	(0.00485)	(0.00604)	(0.00666)	(0.00595)
Children > 0	0.0108*	0.00490	0.000228	-0.0162*	0.000202
	(0.00519)	(0.00552)	(0.00726)	(0.00819)	(0.00724)
2nd Level parent	-0.0302***	0.0261***	-0.00774	0.0136	-0.00183
	(0.00595)	(0.00670)	(0.00872)	(0.0102)	(0.00946)
3rd Level parent	-0.0225***	-0.00992	0.0112	0.00711	0.0141
	(0.00603)	(0.00610)	(0.00830)	(0.00943)	(0.00877)
4th Level parent	-0.0516***	-0.0131*	-0.0160	0.0510***	0.0298***
	(0.00592)	(0.00625)	(0.00830)	(0.00961)	(0.00880)
5th Level parent	-0.0518***	-0.0319***	-0.0268**	0.0472***	0.0633***
	(0.00695)	(0.00690)	(0.00945)	(0.0107)	(0.00976)
Gen. X	0.00510	0.0237***	0.0109	-0.0230**	-0.0167*
	(0.00505)	(0.00491)	(0.00650)	(0.00760)	(0.00668)
Millennial	-0.0131*	0.0413***	0.0364***	-0.0430***	-0.0215**
	(0.00589)	(0.00658)	(0.00814)	(0.00909)	(0.00812)
East = 1	-0.0144	0.0142	-0.0266*	-0.0132	0.0400***
	(0.00786)	(0.00881)	(0.0105)	(0.0124)	(0.0108)

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Delta Method Standard Errors in parentheses. The *Level parent* refers to the highest language skill of (one of) the parents

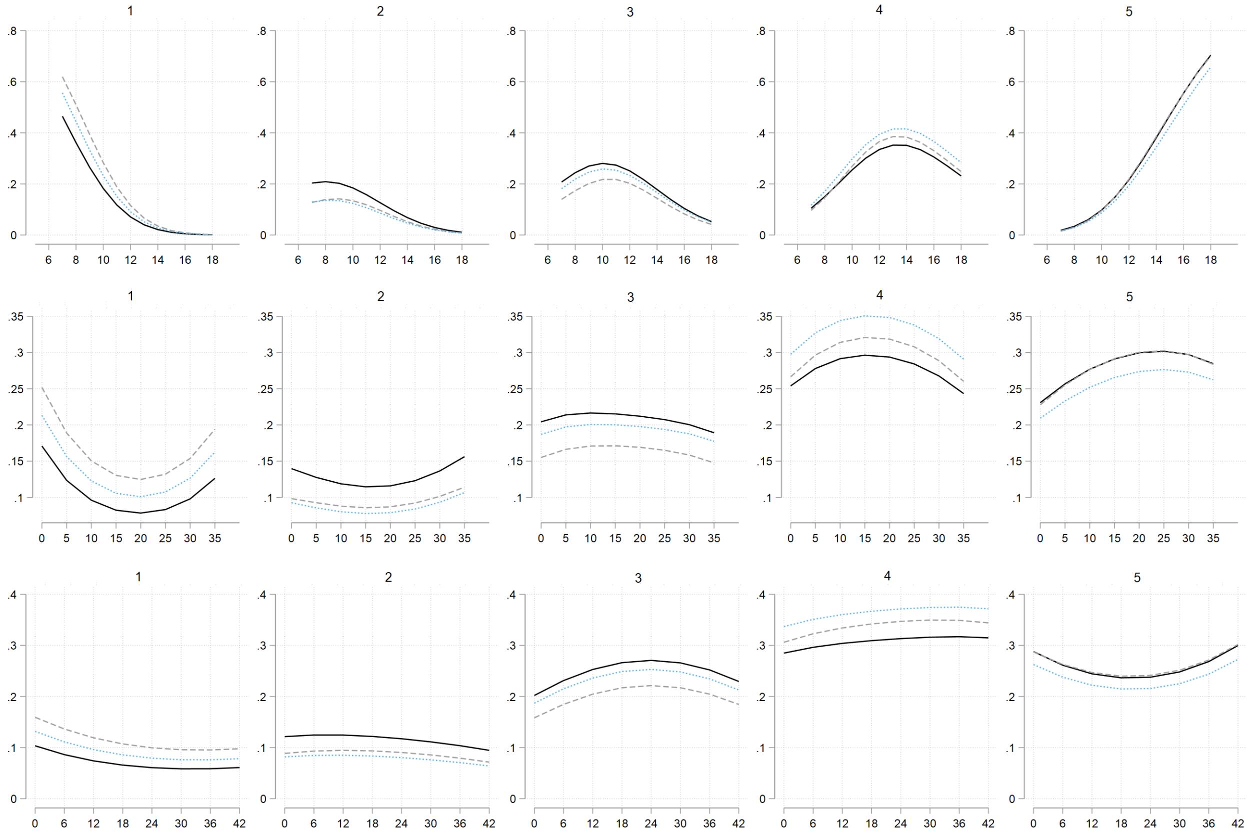


Figure (24) Average Marginal Effects of the continuous variables estimated in the augmented model for Language Skill discriminated by migrant status. (Top) Years of education. (Middle) Years of experience. (Bottom) Years in Germany. In addition, Solid = Native, Dotted = Indirect Migrant, Dashed = Direct Migrant

Table (10) Log-odds of the variations of the Basic Model for ISCO Skill Level.

	(1)				(2)				(3)			
	BLUE-COLLAR	WHITE-COLLAR	TECHNICAL	PROFESSIONAL	BLUE-COLLAR	WHITE-COLLAR	TECHNICAL	PROFESSIONAL	BLUE-COLLAR	WHITE-COLLAR	TECHNICAL	PROFESSIONAL
Female	-2.136***	0.187*	0.134	-0.509***	-2.134***	0.187*	0.135	-0.510***	-2.135***	0.187*	0.135	-0.508***
	(0.085)	(0.083)	(0.078)	(0.081)	(0.085)	(0.083)	(0.078)	(0.081)	(0.085)	(0.083)	(0.078)	(0.081)
Years of education	0.263***	0.512***	0.683***	1.126***	0.264***	0.512***	0.682***	1.124***	0.263***	0.512***	0.683***	1.126***
	(0.029)	(0.029)	(0.029)	(0.029)	(0.029)	(0.029)	(0.028)	(0.029)	(0.029)	(0.029)	(0.029)	(0.029)
Work experience	0.141***	0.073***	0.160***	0.186***	0.141***	0.073***	0.160***	0.185***	0.141***	0.073***	0.160***	0.186***
	(0.017)	(0.016)	(0.016)	(0.017)	(0.017)	(0.016)	(0.016)	(0.017)	(0.017)	(0.016)	(0.016)	(0.017)
Sqr. Work experience	-0.003***	-0.001**	-0.003**	-0.004***	-0.003***	-0.001**	-0.003**	-0.004***	-0.003***	-0.001**	-0.003**	-0.004***
	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)	(0.000)	(0.001)
Western European	0.443	0.368	0.279	0.472	-0.026	0.258	0.103	0.601*	0.393	0.445	0.488	0.901*
	(0.375)	(0.373)	(0.347)	(0.367)	(0.243)	(0.246)	(0.232)	(0.250)	(0.388)	(0.385)	(0.364)	(0.385)
Eastern European	0.277	-0.042	-0.417	-0.942**	-0.187	-0.141	-0.568**	-0.775***	0.208	0.038	-0.186	-0.455
	(0.325)	(0.319)	(0.302)	(0.320)	(0.175)	(0.176)	(0.175)	(0.192)	(0.350)	(0.341)	(0.332)	(0.350)
Turkish/Greek	0.163	0.053	-0.124	-0.771*	-0.282	-0.012	-0.263	-0.651**	0.046	0.139	0.138	-0.263
	(0.335)	(0.333)	(0.314)	(0.337)	(0.183)	(0.184)	(0.182)	(0.208)	(0.370)	(0.362)	(0.352)	(0.376)
Direct migrant	-0.773*	-0.705*	-0.778**	0.093					-1.583*	-0.765	-0.055	1.356*
	(0.323)	(0.319)	(0.300)	(0.318)					(0.805)	(0.745)	(0.658)	(0.657)
Indirect migrant	-0.466	-0.106	-0.206	0.030					-0.386	-0.178	-0.415	-0.358
	(0.324)	(0.317)	(0.295)	(0.308)					(0.336)	(0.327)	(0.312)	(0.327)
Language distance					-0.003	-0.007***	-0.007**	-0.001	0.010	-0.000	-0.011	-0.020*
					(0.002)	(0.002)	(0.002)	(0.002)	(0.009)	(0.009)	(0.008)	(0.008)
N. Observations	22692				22692				22692			
Mc. Fadden R^2	0.216				0.216				0.217			
Adj. Mc. Fadden R^2	0.215				0.215				0.215			
Count	0.532				0.532				0.532			
Adj. Count	0.277				0.277				0.278			
AIC	51746.6				51760.1				51726.9			
Deviance BIC	-13893.5				-13912.1				-13881.1			

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Clustered Standard Errors at the Household level in parentheses.
Baseline level category corresponds to a Male Native German Elementary Worker.

Table (11) Log-odds of the variations of the Basic Model for Language Skill Level.

	(1)				(2)				(3)			
	2	3	4	5	2	3	4	5	2	3	4	5
Female	-0.079	0.720***	1.036***	0.986***	-0.077	0.721***	1.036***	0.988***	-0.078	0.721***	1.037***	0.986***
	(0.065)	(0.057)	(0.056)	(0.059)	(0.065)	(0.057)	(0.056)	(0.059)	(0.065)	(0.057)	(0.056)	(0.059)
Years of education	0.284***	0.449***	0.675***	0.942***	0.283***	0.448***	0.674***	0.940***	0.284***	0.448***	0.674***	0.942***
	(0.022)	(0.020)	(0.020)	(0.020)	(0.022)	(0.020)	(0.020)	(0.020)	(0.022)	(0.020)	(0.020)	(0.020)
Work experience	0.046**	0.089***	0.119***	0.134***	0.046**	0.089***	0.119***	0.135***	0.046**	0.089***	0.119***	0.134***
	(0.014)	(0.013)	(0.013)	(0.014)	(0.014)	(0.013)	(0.013)	(0.014)	(0.014)	(0.013)	(0.013)	(0.014)
Sqr. Work experience	-0.001**	-0.003***	-0.003***	-0.003***	-0.001**	-0.002***	-0.003***	-0.003***	-0.001**	-0.002***	-0.003***	-0.003***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Western European	0.692*	0.457	0.098	0.575	0.238	0.438*	0.297	0.548**	0.744*	0.616*	0.367	1.011**
	(0.348)	(0.291)	(0.283)	(0.313)	(0.210)	(0.186)	(0.188)	(0.206)	(0.360)	(0.301)	(0.298)	(0.324)
Eastern European	0.365	-0.205	-0.662**	-0.506	-0.076	-0.208	-0.439**	-0.486**	0.428	-0.021	-0.353	0.006
	(0.308)	(0.255)	(0.247)	(0.278)	(0.145)	(0.133)	(0.140)	(0.154)	(0.329)	(0.272)	(0.271)	(0.297)
Turkish/Greek	0.501	0.199	-0.351	-0.161	0.057	0.208	-0.130	-0.195	0.581	0.431	0.019	0.401
	(0.316)	(0.263)	(0.261)	(0.294)	(0.151)	(0.138)	(0.153)	(0.176)	(0.346)	(0.287)	(0.292)	(0.321)
Direct migrant	-0.564	-0.407	-0.130	-0.247					-0.331	0.321	1.034*	1.576**
	(0.306)	(0.253)	(0.245)	(0.277)					(0.701)	(0.541)	(0.506)	(0.524)
Indirect migrant	-0.489	-0.043	0.160	-0.140					-0.544	-0.220	-0.124	-0.567*
	(0.307)	(0.249)	(0.242)	(0.262)					(0.316)	(0.256)	(0.254)	(0.273)
Language distance					-0.001	-0.005**	-0.004**	-0.003	-0.003	-0.010	-0.017**	-0.027***
					(0.002)	(0.001)	(0.002)	(0.002)	(0.008)	(0.006)	(0.006)	(0.007)
N. Observations	22704				22704				22704			
Mc. Fadden R^2	0.128				0.128				0.129			
Adj. Mc. Fadden R^2	0.127				0.127				0.127			
Count	0.404				0.404				0.404			
Adj. Count	0.157				0.157				0.157			
AIC	60466.1				60465.6				60451.1			
Deviance BIC	-8537.6				-8570.2				-8520.5			

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Clustered Standard Errors at the Household level in parentheses.
Baseline level category corresponds to a Male Native German 1st Level Worker.

Table (12) Log-odds: Second generation - Native for ISCO Skill Level.

	(Basic)				(Augmented)			
	BLUE-COLLAR	WHITE-COLLAR	TECHNICAL	PROFESSIONAL	BLUE-COLLAR	WHITE-COLLAR	TECHNICAL	PROFESSIONAL
Female	-2.100***	0.391***	0.301***	-0.352***	-2.095***	0.486***	0.350***	-0.285***
	(0.076)	(0.070)	(0.067)	(0.073)	(0.083)	(0.078)	(0.075)	(0.081)
Years of education	0.262***	0.550***	0.737***	1.160***	0.227***	0.533***	0.715***	1.129***
	(0.029)	(0.029)	(0.029)	(0.030)	(0.035)	(0.034)	(0.034)	(0.035)
Work experience	0.128***	0.090***	0.136***	0.139***	0.146***	0.113***	0.158***	0.161***
	(0.016)	(0.015)	(0.015)	(0.017)	(0.018)	(0.017)	(0.016)	(0.018)
Sqr. Work experience	-0.002***	-0.002***	-0.002***	-0.002***	-0.003***	-0.002***	-0.003***	-0.003***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.001)
Germanic	-0.008	0.097	0.057	0.404	-0.079	-0.192	-0.087	0.282
	(0.332)	(0.318)	(0.307)	(0.318)	(0.354)	(0.339)	(0.327)	(0.339)
Western european	0.288	0.149	0.130	-0.011	0.423	-0.066	-0.023	-0.156
	(0.271)	(0.283)	(0.290)	(0.359)	(0.305)	(0.312)	(0.335)	(0.416)
Eastern european	-0.320	-0.035	-0.267	-0.142	-0.336	-0.209	-0.346	-0.113
	(0.214)	(0.203)	(0.205)	(0.228)	(0.234)	(0.224)	(0.229)	(0.252)
Turkish/Greek	-0.138	0.151	-0.364*	-0.462*	-0.127	-0.132	-0.471*	-0.323
	(0.164)	(0.157)	(0.171)	(0.216)	(0.186)	(0.183)	(0.193)	(0.238)
Education East					0.080	-0.580***	-0.750***	-0.839***
					(0.082)	(0.080)	(0.079)	(0.086)
Blue-collar Parent					0.654***	0.280*	0.322*	0.316
					(0.144)	(0.139)	(0.136)	(0.174)
White-collar Parent					0.379*	0.454**	0.476**	0.458*
					(0.163)	(0.156)	(0.154)	(0.191)
Technician Parent					0.600***	0.604***	0.820***	0.859***
					(0.154)	(0.147)	(0.144)	(0.180)
Professional Parent					0.482**	0.579***	0.850***	1.167***
					(0.172)	(0.164)	(0.161)	(0.193)
Gen. X					0.023	0.291***	-0.119	-0.430***
					(0.089)	(0.087)	(0.084)	(0.091)
Millenial					0.295**	0.596***	0.096	-0.227*
					(0.106)	(0.102)	(0.099)	(0.108)
N. Observations	23517				20969			
Adj. Mc. Fadden R^2	0.209				0.222			
Adj. Count	0.0216				0.0207			
AIC	2.387				2.336			
Deviance BIC	-14575.1				-13555.9			

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Clustered Standard Errors at the Household level in parentheses.
Baseline level category corresponds to a Male Native German Elementary Worker.

Table (13) Log-odds of the variations of the Augmented Model for ISCO Skill Level.

	(1)				(2)				(3)			
	BLUE-COLLAR	WHITE-COLLAR	TECHNICAL	PROFESSIONAL	BLUE-COLLAR	WHITE-COLLAR	TECHNICAL	PROFESSIONAL	BLUE-COLLAR	WHITE-COLLAR	TECHNICAL	PROFESSIONAL
Female	-2.091***	0.267**	0.174*	-0.481***	-2.091***	0.268**	0.174*	-0.482***	-2.091***	0.267**	0.174*	-0.481***
	(0.086)	(0.084)	(0.079)	(0.082)	(0.086)	(0.084)	(0.079)	(0.082)	(0.086)	(0.084)	(0.079)	(0.082)
Years of education	0.232***	0.481***	0.648***	1.072***	0.232***	0.481***	0.648***	1.071***	0.233***	0.482***	0.648***	1.073***
	(0.031)	(0.031)	(0.031)	(0.031)	(0.031)	(0.031)	(0.031)	(0.031)	(0.031)	(0.031)	(0.031)	(0.031)
Work experience	0.148***	0.088***	0.163***	0.187***	0.147***	0.088***	0.162***	0.186***	0.148***	0.088***	0.163***	0.187***
	(0.018)	(0.017)	(0.016)	(0.018)	(0.018)	(0.017)	(0.016)	(0.018)	(0.018)	(0.017)	(0.016)	(0.018)
Sqr. Work experience	-0.002***	-0.001**	-0.003***	-0.004***	-0.002***	-0.001**	-0.003***	-0.004***	-0.002***	-0.001**	-0.003***	-0.004***
	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)	(0.000)	(0.001)
Western European	0.506	0.464	0.322	0.576	0.016	0.203	-0.033	0.589*	0.477	0.550	0.536	0.996*
	(0.386)	(0.384)	(0.359)	(0.375)	(0.266)	(0.269)	(0.258)	(0.273)	(0.397)	(0.393)	(0.374)	(0.392)
Eastern European	0.271	-0.050	-0.438	-0.782*	-0.213	-0.295	-0.760***	-0.740***	0.228	0.040	-0.201	-0.310
	(0.338)	(0.333)	(0.317)	(0.331)	(0.204)	(0.205)	(0.208)	(0.221)	(0.359)	(0.351)	(0.342)	(0.358)
Turkish/Greek	0.238	0.161	-0.024	-0.461	-0.215	-0.045	-0.316	-0.424	0.151	0.261	0.248	0.035
	(0.349)	(0.346)	(0.328)	(0.348)	(0.214)	(0.215)	(0.214)	(0.237)	(0.378)	(0.371)	(0.361)	(0.382)
Education east	0.019	-0.314	-0.629***	-0.866***	0.031	-0.306	-0.619***	-0.863***	0.019	-0.315*	-0.630***	-0.869***
	(0.161)	(0.160)	(0.153)	(0.159)	(0.161)	(0.160)	(0.153)	(0.159)	(0.161)	(0.160)	(0.153)	(0.159)
Education abroad	-0.396*	-0.638***	-0.546***	-0.216	-0.470**	-0.687***	-0.593***	-0.117	-0.391*	-0.641***	-0.560***	-0.239
	(0.164)	(0.165)	(0.159)	(0.170)	(0.160)	(0.161)	(0.156)	(0.168)	(0.164)	(0.165)	(0.160)	(0.170)
Years in Germany	0.089***	0.090***	0.118***	0.061*	0.081***	0.085***	0.114***	0.085***	0.089***	0.090***	0.118***	0.061*
	(0.023)	(0.024)	(0.023)	(0.025)	(0.023)	(0.023)	(0.023)	(0.025)	(0.023)	(0.024)	(0.023)	(0.025)
Sqr. Years in Germany	-0.002**	-0.002**	-0.002***	-0.001	-0.002**	-0.002**	-0.002***	-0.002**	-0.002**	-0.002***	-0.002***	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Married to German	0.144	0.157	0.318***	0.423***	0.153	0.162	0.325***	0.425***	0.146	0.157	0.317***	0.419***
	(0.092)	(0.091)	(0.086)	(0.090)	(0.092)	(0.091)	(0.086)	(0.090)	(0.092)	(0.091)	(0.086)	(0.090)
Children > 0	-0.104	-0.188	-0.192	-0.160	-0.099	-0.183	-0.186	-0.157	-0.104	-0.186	-0.188	-0.152
	(0.111)	(0.110)	(0.105)	(0.110)	(0.110)	(0.110)	(0.105)	(0.110)	(0.111)	(0.110)	(0.105)	(0.110)
Blue-collar parent	0.517***	0.172	0.452***	0.477***	0.519***	0.174	0.454***	0.480***	0.519***	0.173	0.454***	0.479***
	(0.124)	(0.127)	(0.123)	(0.141)	(0.124)	(0.127)	(0.123)	(0.140)	(0.125)	(0.127)	(0.123)	(0.141)
White-collar parent	0.222	0.349*	0.459**	0.507**	0.222	0.349*	0.459**	0.510**	0.223	0.350*	0.460**	0.510**
	(0.146)	(0.145)	(0.141)	(0.160)	(0.146)	(0.145)	(0.140)	(0.159)	(0.146)	(0.145)	(0.141)	(0.160)
Technician parent	0.607***	0.548***	1.010***	1.122***	0.611***	0.551***	1.013***	1.126***	0.607***	0.548***	1.010***	1.125***
	(0.142)	(0.142)	(0.137)	(0.154)	(0.142)	(0.142)	(0.137)	(0.154)	(0.142)	(0.142)	(0.137)	(0.154)

Table 13 Continued

	(1)				(2)				(3)			
	BLUE-COLLAR	WHITE-COLLAR	TECHNICAL	PROFESSIONAL	BLUE-COLLAR	WHITE-COLLAR	TECHNICAL	PROFESSIONAL	BLUE-COLLAR	WHITE-COLLAR	TECHNICAL	PROFESSIONAL
Professional parent	0.563***	0.670***	1.129***	1.518***	0.564***	0.670***	1.130***	1.524***	0.564***	0.668***	1.127***	1.518***
	(0.169)	(0.167)	(0.161)	(0.175)	(0.169)	(0.167)	(0.160)	(0.174)	(0.169)	(0.167)	(0.160)	(0.175)
Gen. X	0.127	0.148	-0.054	-0.296**	0.121	0.144	-0.059	-0.301**	0.128	0.149	-0.053	-0.294**
	(0.099)	(0.102)	(0.095)	(0.099)	(0.099)	(0.102)	(0.095)	(0.099)	(0.099)	(0.102)	(0.095)	(0.099)
Millennial	0.390***	0.475***	0.028	-0.246*	0.367**	0.458***	0.006	-0.250*	0.391***	0.474***	0.026	-0.247*
	(0.117)	(0.116)	(0.110)	(0.115)	(0.116)	(0.115)	(0.109)	(0.114)	(0.117)	(0.116)	(0.110)	(0.115)
East	0.329*	-0.014	0.154	0.435**	0.330*	-0.012	0.158	0.439**	0.330*	-0.012	0.159	0.442**
	(0.153)	(0.153)	(0.147)	(0.151)	(0.153)	(0.153)	(0.147)	(0.152)	(0.153)	(0.153)	(0.147)	(0.151)
Direct migrant	-1.114**	-1.017*	-1.377***	-0.362					-1.861*	-1.017	-0.580	0.936
	(0.403)	(0.398)	(0.375)	(0.400)					(0.862)	(0.812)	(0.713)	(0.716)
Indirect migrant	-0.532	-0.285	-0.425	-0.027					-0.474	-0.367	-0.639*	-0.409
	(0.337)	(0.329)	(0.308)	(0.320)					(0.345)	(0.334)	(0.321)	(0.334)
Language distance					-0.005	-0.008**	-0.011***	-0.008*	0.009	-0.001	-0.012	-0.020*
					(0.003)	(0.003)	(0.003)	(0.003)	(0.010)	(0.009)	(0.008)	(0.008)
N. Observations	22692				22692				22692			
Mc. Fadden R^2	0.229				0.229				0.229			
Adj. Mc. Fadden R^2	0.226				0.226				0.226			
Count	0.543				0.542				0.543			
Adj. Count	0.295				0.293				0.295			
AIC	51012.8				51019.2				50996.5			
Deviance BIC	-14209.8				-14235.5				-14194.0			

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Clustered Standard Errors at the Household level in parentheses.
Baseline level category corresponds to a single baby boomer native german male working in an elementary occupation, educated in West Germany, with no Children and an elementary worker parental background, living in West Germany.

Table (14) Log-odds of the variations of the Augmented Model for Language Skill Level.

	(1)				(2)				(3)			
	2	3	4	5	2	3	4	5	2	3	4	5
Female	0.008	0.787***	1.085***	1.045***	0.009	0.787***	1.085***	1.046***	0.008	0.787***	1.085***	1.044***
	(0.066)	(0.058)	(0.057)	(0.061)	(0.066)	(0.058)	(0.057)	(0.061)	(0.066)	(0.058)	(0.057)	(0.061)
Years of education	0.285***	0.445***	0.655***	0.914***	0.283***	0.444***	0.654***	0.913***	0.284***	0.445***	0.655***	0.914***
	(0.023)	(0.022)	(0.022)	(0.022)	(0.023)	(0.022)	(0.022)	(0.022)	(0.023)	(0.022)	(0.022)	(0.022)
Work experience	0.059***	0.102***	0.124***	0.140***	0.059***	0.102***	0.124***	0.140***	0.059***	0.102***	0.124***	0.140***
	(0.014)	(0.013)	(0.013)	(0.014)	(0.014)	(0.013)	(0.013)	(0.014)	(0.014)	(0.013)	(0.013)	(0.014)
Sqr. Work experience	-0.001**	-0.003***	-0.003***	-0.003***	-0.001**	-0.003***	-0.003***	-0.003***	-0.001**	-0.003***	-0.003***	-0.003***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Western European	0.793*	0.526	0.197	0.704*	0.222	0.263	0.153	0.475*	0.858*	0.690*	0.464	1.138***
	(0.351)	(0.296)	(0.288)	(0.317)	(0.222)	(0.202)	(0.204)	(0.222)	(0.361)	(0.305)	(0.302)	(0.328)
Eastern European	0.369	-0.221	-0.577*	-0.341	-0.188	-0.464**	-0.598***	-0.532**	0.447	-0.031	-0.270	0.164
	(0.312)	(0.261)	(0.253)	(0.283)	(0.164)	(0.156)	(0.161)	(0.177)	(0.330)	(0.277)	(0.276)	(0.302)
Turkish/Greek	0.568	0.241	-0.198	0.077	0.011	0.009	-0.215	-0.145	0.667	0.482	0.172	0.634
	(0.320)	(0.270)	(0.267)	(0.300)	(0.172)	(0.162)	(0.174)	(0.197)	(0.347)	(0.291)	(0.296)	(0.325)
Education east	-0.399**	-0.587***	-0.871***	-1.028***	-0.388**	-0.581***	-0.868***	-1.025***	-0.399**	-0.589***	-0.873***	-1.031***
	(0.125)	(0.115)	(0.114)	(0.119)	(0.126)	(0.116)	(0.114)	(0.119)	(0.125)	(0.115)	(0.114)	(0.119)
Education abroad	-0.326*	-0.492***	-0.418**	-0.225	-0.360**	-0.509***	-0.380**	-0.165	-0.333*	-0.504***	-0.439***	-0.249
	(0.132)	(0.133)	(0.130)	(0.138)	(0.127)	(0.127)	(0.125)	(0.135)	(0.132)	(0.133)	(0.131)	(0.139)
Years in Germany	0.040*	0.060**	0.040*	0.009	0.038*	0.061***	0.048**	0.028	0.041*	0.061***	0.040*	0.010
	(0.020)	(0.018)	(0.019)	(0.021)	(0.019)	(0.018)	(0.018)	(0.021)	(0.020)	(0.018)	(0.019)	(0.021)
Sqr. Years in Germany	-0.001	-0.001*	-0.001	0.000	-0.001	-0.001*	-0.001	-0.000	-0.001	-0.001*	-0.001	0.000
	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)
Married to German	0.085	0.180**	0.194**	0.267***	0.094	0.185**	0.195**	0.271***	0.084	0.178**	0.190**	0.261***
	(0.072)	(0.063)	(0.062)	(0.066)	(0.072)	(0.063)	(0.062)	(0.066)	(0.072)	(0.063)	(0.062)	(0.066)
Children > 0	-0.070	-0.126	-0.192**	-0.146	-0.066	-0.123	-0.190**	-0.142	-0.068	-0.122	-0.186*	-0.137
	(0.083)	(0.074)	(0.073)	(0.078)	(0.083)	(0.074)	(0.073)	(0.078)	(0.083)	(0.074)	(0.073)	(0.078)
Blue-collar parent	0.470***	0.238**	0.331***	0.283**	0.471***	0.239**	0.333***	0.286***	0.471***	0.239**	0.333***	0.286***
	(0.082)	(0.077)	(0.077)	(0.086)	(0.082)	(0.077)	(0.077)	(0.086)	(0.082)	(0.077)	(0.077)	(0.087)
White-collar parent	0.111	0.274***	0.269***	0.321***	0.111	0.273***	0.269***	0.324***	0.111	0.273***	0.269***	0.324***
	(0.084)	(0.074)	(0.074)	(0.082)	(0.084)	(0.074)	(0.074)	(0.082)	(0.084)	(0.074)	(0.074)	(0.082)
Technician parent	0.408***	0.503***	0.805***	0.806***	0.410***	0.505***	0.807***	0.811***	0.408***	0.503***	0.806***	0.810***
	(0.091)	(0.082)	(0.081)	(0.088)	(0.091)	(0.082)	(0.081)	(0.088)	(0.091)	(0.082)	(0.081)	(0.088)

Table 14 Continued

	(1)				(2)				(3)			
	2	3	4	5	2	3	4	5	2	3	4	5
Professional parent	0.219	0.477***	0.846***	1.003***	0.220	0.478***	0.848***	1.009***	0.218	0.475***	0.844***	1.004***
	(0.115)	(0.101)	(0.098)	(0.104)	(0.115)	(0.101)	(0.098)	(0.104)	(0.115)	(0.101)	(0.098)	(0.104)
Gen. X	0.173*	-0.028	-0.186**	-0.201**	0.168*	-0.032	-0.189**	-0.206**	0.173*	-0.028	-0.186**	-0.199**
	(0.076)	(0.068)	(0.066)	(0.071)	(0.076)	(0.068)	(0.066)	(0.070)	(0.076)	(0.068)	(0.066)	(0.071)
Millenial	0.495***	0.266**	-0.090	-0.062	0.477***	0.254**	-0.093	-0.074	0.493***	0.263**	-0.093	-0.064
	(0.094)	(0.083)	(0.082)	(0.088)	(0.094)	(0.083)	(0.082)	(0.087)	(0.094)	(0.083)	(0.082)	(0.088)
East	0.279*	0.033	0.148	0.373**	0.279*	0.034	0.150	0.376**	0.281*	0.036	0.153	0.379**
	(0.122)	(0.114)	(0.112)	(0.117)	(0.122)	(0.114)	(0.112)	(0.117)	(0.122)	(0.114)	(0.112)	(0.117)
Direct migrant	-0.770*	-0.712*	-0.403	-0.478					-0.454	0.111	0.837	1.399*
	(0.360)	(0.308)	(0.302)	(0.338)					(0.757)	(0.602)	(0.572)	(0.596)
Indirect migrant	-0.649*	-0.323	-0.078	-0.350					-0.718*	-0.505	-0.359	-0.774**
	(0.312)	(0.257)	(0.250)	(0.270)					(0.318)	(0.262)	(0.259)	(0.279)
Language distance					-0.002	-0.005*	-0.005*	-0.005*	-0.004	-0.011	-0.017**	-0.027***
					(0.002)	(0.002)	(0.002)	(0.003)	(0.008)	(0.007)	(0.007)	(0.007)
N. Observations	22704				22704				22704			
Mc. Fadden R^2	0.138				0.138				0.138			
Adj. Mc. Fadden R^2	0.135				0.135				0.135			
Count	0.416				0.416				0.415			
Adj. Count	0.295				0.293				0.295			
AIC	59935.5				59933.1				59921.7			
Deviance BIC	-8650.6				-8685.1				-8632.3			

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Clustered Standard Errors at the Household level in parentheses.
Baseline level category corresponds to a single baby boomer native german male working in an elementary occupation, educated in West Germany, with no Children and an elementary worker parental background, living in West Germany.

C. STATISTICAL TESTS

In the following subsection, we revisit the ideas behind the statistical tests performed to assess the estimations and briefly mention the general idea behind their results. However, the set of full estimations can be handed anytime upon request.

C.1. Independent Variables and Combinations of Dependent Variable Categories

Following Train (2009), since the Multinomial Logistic Regression is solved using a Maximum-Likelihood estimator, it is possible to assess the Goodness of fit through a *Likelihood-Ratio* Test. Basically, it tests the null hypothesis of an arbitrary number of parameters being jointly zero or equal between themselves. To generate the comparison statistic, a constrained model is estimated in which the null hypothesis is binding. If $\hat{\beta}^H$ are the estimates of the constrained model that maximize the value of the likelihood function $L(.)$ and $\hat{\beta}$ the ones of the unconstrained version, the ratio is defined as $R = L(\hat{\beta}^H)/L(\hat{\beta})$. The test statistic defined as $-2 \times log(R) = -2 \times [log(L(\hat{\beta}^H)) - log(L(\hat{\beta}^H)]$ is distributed chi-squared with degrees of freedom equal to the number of restrictions implied by the null hypothesis.

In the other hand, under certain regularity conditions, the Maximum-Likelihood estimator distributes approximately normal with mean equal to the true parameter and variance-covariance matrix given by the inverse of the information matrix, therefore, the statistical significance of individual variables could also be checked by performing a Wald test. If we define the standard error of β as the squared root of the diagonal element of its corresponding row-column position on the variance-covariance matrix. The Wald statistic for $\hat{\beta} = 0$ would be:

$$\hat{z} = \frac{\hat{\beta}_i}{\sqrt{\hat{Var}(\hat{\beta}_i)}} \sim Z \text{ with one degree of freedom}$$

Finally, according to Freese and Long (2000) both tests could be generalized to check whether is possible to merge two categories within the regressed dependent variable. The basic idea relies on the fact that two categories are *indistinguishable* by the model if all their corresponding estimates are equal. Then, if there are p number of estimated parameters, and i and j are two different categories while $base$ is the baseline category, the null hypothesis would be:

$$H_o : (\beta_{1,i|base} - \beta_{1,j|base}) = ... = (\beta_{p,i|base} - \beta_{p,j|base}) = 0$$

For the augmented models, including both ISCO and Language Skill, we find that having at least one children is not significant in any variation. Furthermore, even when at the descriptive level cultural differences are evident, culture exhibit no clear pattern of significance between models and skill measures, yielding statistical significance for some groups, while large standard errors for others, therefore, cultural clusters should remain under close inspection, especially if one is able to measure work-related values heterogeneity within cultures. Finally, it appears that language distance and migration background are capturing the same phenomenon, and when included at the same time, migration background lose its significance —especially the indirect migration dummy—, probably due to the fact that it is capturing less heterogeneity compared to the language distance counterpart. In general, since we are interested in modeling Native-Immigrant Gaps, we chose those models with migration backgrounds instead of language distances, however the language distance remains as an interesting variable to further explorations.

C.2. Independence of Irrelevant Alternatives (IIA)

According to Train (2009), testing the IIA condition consists in estimating two models, one full and other restricting one of the alternatives, and then comparing both estimated coefficients with a statistical test. If the test yields statistically significant results, then IIA is violated. Note, however, that in general IIA cannot be proved nor disproved completely by an statistical argument, since the condition imposes restrictions on the decision-maker behavior and rationalization processes, thus, implicitly imposing strong assumptions to even identify the model. At that respect, Cheng and Long (2007) ran Monte Carlo experiments to assess the properties of IIA tests, still they found that the sample size properties are very poor and do not improve asymptotically. Therefore, as showed by Long and Freese (2014), since Hausmann-McFadden and Small-Hsiao show erratic and inconsistent behaviors, we won't report the IIA tests in the current document. Still, results are available upon request.

D. OCCUPATIONAL CLASSIFICATIONS COMPARED

ISCO-08 Code	Name of the job unit
UNCHANGED	
Elementary	
9111	Charworker, domestic
9112	Attendant, lavatory
9121	Cleaner, dry: hand
9122	Cleaner, vehicles
9123	Cleaner, window
9129	Blaster, water: cleaning
9211	Cutter, sugar cane
9212	Collector, egg
9213	Hand, farm
9214	Assistant, gardener's
9215	Axeman
9216	Beachcomber
9311	Labourer, mining
9312	Digger, grave: hand-held tools
9313	Assistant, bricklayer's
9321	Bagger, hand
9329	Cellarhand, wine production
9331	Boy, rickshaw
9332	Driver, animal
9333	Attendant, airport: handling baggage
9334	Assistant, cabinet: supermarket
9611	Collector, garbage
9612	Dealer, scrap
9613	Cleaner, park
9624	Bhishti
Blue-collar	
6111	Farmer, alfalfa
6112	Arboriculturist
6113	Contractor, gardening
6114	Farmer, mixed crop
6121	Breeder, cat
6122	Breeder, poultry
6123	Apiarist
6129	Breeder, bird
6130	Farmer, livestock and crops
6210	Assembler, raft: logging
6221	Cultivator, algae
6222	Captain, fishing: coastal waters
6223	Captain, fishing: deep-sea
6224	Beater, game
6320	Cowherd: subsistence farming
7114	Caster, concrete: products
7119	Assembler, prefabricated building
7121	Fixer, roof
7122	Blocklayer, wood
7124	Installer, insulation
7131	Hanger, wallpaper
7132	Painter, automobile
7133	Blaster sand, building exteriors
7213	Beater, aircraft panel
7214	Bender, metal plate
7215	Grip
7221	Blacksmith
7222	Diemaker
7223	Borer, metal
7224	Bluer, metal
7231	Fitter, diesel: road transport
7233	Erector-installer, agricultural machinery
7412	Builder, armature
7413	Jointer, cable: electric
7512	Baker
7515	Grader, food
7521	Dipper, wood treatment
7541	Diver, abalone
7542	Blaster
7544	Contractor, spraying: pest or weed control
7549	Arranger, flower
8121	Annealer
8122	Anodiser
8131	Candle-maker, machine
8142	Builder, plastic boat
8143	Moulder, papier maché
8182	Engineer, stationary
8183	Filler, bottle
8211	Assembler, agricultural machinery
8212	Assembler, armature
8219	Assembler, ammunition
8350	Boatman
White-collar	
4110	Clerk
4211	Assistant, bank
4222	Agent, directory assistance
4227	Assistant, survey: interviewing
4312	Assistant, broker's
4322	Clerk, order: materials
4411	Assistant, library
4414	Clerk, form filling: assistance
4415	Clerk, copying
4419	Clerk, addressing machine
5111	Attendant, cabin
5112	Attendant, pullman car
5113	Director, tour
5120	Cook
5131	Attendant, restaurant seating
5132	Attendant, bar: drinks service
5141	Barber
5142	Artist, make-up
5161	Astrologer
5162	Companion, except health or aged care
5163	Attendant, funeral
5164	Aide, veterinary
5165	Examiner, driving
5169	Escort, social
5221	Florist, operating a shop
5222	Supervisor, checkout
5223	Agent, leasing: vehicle
5242	Demonstrator, sales
5244	Consultant, sales: outbound calls
Technical	
1431	Director, cultural centre
1439	Manager, camp site
3114	Assistant, computer: engineering (hardware)
3211	Mammographer
3212	Technician, blood-bank
3214	Denturist
3221	Nurse, assistant
3222	Midwife, assistant
3240	Assistant, veterinary
3253	Aide, community health
3312	Analyst, credit: assessing credit or loans
3321	Agent, group insurance
3322	Adviser, after-sales service
3323	Agent, procurement
3341	Administrator, office
3342	Assistant, administrative: legal
3343	Assistant, administrative
3354	Inspector, licensing
3411	Agent, inquiry: private
3412	Almoner, associate professional
3413	Brother
3432	Decorator, display
3511	Assistant, computer: engineering (operations)
3512	Agent, technical support: information technology
3513	Assistant , communications: ICT
3514	Administrator, website
3521	Assistant, production: media
3522	Officer, ship: radio
Professional	
1212	Director, human resources
1221	Director, marketing
1222	Director, advertising
1223	Director, business development: except ICT
1342	Administrator, health facility
1344	Manager, centre: welfare services
2111	Aerodynamicist
2112	Climatologist
2113	Chemist
2114	Gemmologist
2120	Actuary
2131	Anatomist
2132	Adviser, agricultural
2133	Adviser, environmental
2141	Engineer, food processing
2142	Engineer, building structure
2143	Analyst, environmental
2149	Analyst, systems: except computers
2162	Architect, landscape
2163	Designer, clothing
2164	Planner, land
2166	Animator
2221	Anaesthetist, nurse
2222	Educator, midwife
2240	Assistant, clinical: diagnosing and treating patients
2250	Epidemiologist, veterinary
2265	Consultant, dietetic
2266	Audiologist
2310	Academic, university: lecturer
2320	Instructor, automotive technology
2330	Master, high school
2341	Master, primary education
2342	Educator, early childhood
2351	Adviser, academic
2352	Teacher, for the blind
2353	Teacher, EFL
2355	Coach, dance
2356	Consultant, computer training
2359	Adviser, student
2421	Analyst, cost
2422	Adviser, political
2423	Adviser, careers
2424	Assessor, training
2431	Analyst, market: research
2432	Agent, publicity
2511	Administrator, SAP: business analysis
2512	Analyst, programme: computers
2513	Architect, information: computing (website)
2514	Programmer, applications
2519	Analyst, business: testing software
2521	Administrator, data
2522	Administrator, computer systems
2523	Analyst, communications: computers
2529	Analyst, data mining
2611	Adviser, legal
2612	Chief, justice
2619	Coroner
2621	Archivist
2622	Bibliographer
2631	Adviser, economic
2632	Anthropologist
2633	Futurologist
2634	Analyst, psychological
2635	Almoner, professional
2636	Abbess
2641	Author
2642	Blogger
2643	Etymologist
2654	Director, artistic
2655	Actor
CHANGED	
Elementary to 2nd Language Level	
9411	Boy, pizza: maker
9412	Assistant, kitchen
9510	Boy, errand
9520	Hawker, except food
9621	Attendant, lift
9622	Attendant, cellar: hotel
9623	Collector, coin machine
9629	Attendant, amusement park
Elementary to 3th Language Level	
310	Aircrew woman
Blue-collar to 1st Language Level	
6310	Farmer, cereal: subsistence farming
6330	Farmer, mixed: subsistence
6340	Collector, subsistence
7112	Blocklayer
7113	Blaster, sand: stonecutting
7123	Fixer, plasterboard
7125	Autoglazier
7126	Fitter, aircraft pipe
7211	Coremaker
7212	Brazier
7311	Adjuster, precision instrument
7313	Beater, gold
7314	Burnisher, ceramics
7315	Bender, glass
7316	Calligrapher
7317	Creeler
7318	Bleacher, fibre: textile
7319	Candle-maker, handicraft
7322	Cutter, stencil: silk-screen
7323	Binder, book
7511	Boner, fish
7513	Maker, butter
7514	Brewer, not operating machinery
7516	Blender, snuff
7522	Applier, veneer
7523	Borer, wood
7532	Copyist, jacquard design
7533	Embroiderer
7534	Maker, bedding
7535	Buffer, leather
7536	Burnisher, footwears
8111	Bolter, roof: mining
8112	Operator, breaker: gyratory
8113	Borer, well
8114	Cutter-polisher, industrial diamonds
8132	Developer, film: black-and-white
8141	Maker, tyre
8151	Baller, thread and yarn
8152	Cutter, jacquard card
8153	Machinist, sewing
8154	Calenderer, cloth
8155	Operator, machine: cutting (leather)
8156	Operator, machine: footwear production
8157	Calenderer, laundry
8159	Blocker, hat
8160	Brewer, operating machinery

ISCO-08 Code	Name of the job unit
8171	Calenderer, pulp and paper
8172	Assembler, plywood panel
8181	Operator, die-press: pottery and porcelain
8311	Driver, assistant: railway-engine
8312	Braker, railway
8321	Courier, motorcycle
8322	Attendant, car park: driving cars
8331	Driver, bus
8332	Driver, aircraft fueller
8341	Driver, lumber carrier
8342	Digger, grave: earthmoving equipment
8343	Attendant, dry dock
8344	Driver, truck: forklift
Blue-collar to 3th Language Level	
7115	Boatbuilder, wood
7127	Erector, refrigeration and air conditioning equipment
7232	Aeromechanic
7234	Mechanic, bicycle
7312	Builder, organ
7411	Electrician
7422	Cabler, data
7531	Coner, hat forms
7543	Classer, wool
8189	Operator, machine: pencil production
Blue-collar to 4th Language Level	
7111	Builder, house
7321	Cameraman, photogravure
7421	Engineer, aircraft maintenance: avionics
Blue-collar to 5th Language Level	
210	Airman, air force: warrant officer
White-collar to 1st Language Level	
5153	Caretaker, building
5230	Agent, ticket: entertainment and sporting events
5241	Mannequin
White-collar to 2nd Language Level	
4311	Clerk, accounts
4412	Carrier, post
4416	Assistant, human resource
5152	Butler
5211	Assistant, sales: market stall
5212	Hawker, food
5243	Canvasser, door-to-door
5245	Attendant, driveway
5246	Attendant, bar: food service
5249	Assistant, sales: car hire
5311	Assistant, day care: children
5312	Aide, pre-school
5321	Aide, nursing: clinic
5322	Aide, home care
5329	Aide, dental
White-collar to 4th Language Level	
4212	Bookmaker
4213	Lender, money
4221	Adviser, travel
4223	Operator, answering service
4225	Clerk, enquiry
4226	Clerk, appointments
4229	Clerk, hospital admissions
4313	Clerk, payroll
4321	Attendant, tool crib
5151	Housekeeper, executive
5411	Fighter, fire
5413	Gaoler
5414	Bodyguard
5419	Attendant, pool
White-collar to 5th Language Level	
4120	Secretary
4131	Clerk, justowriting
4132	Clerk, accounting machine
4214	Clerk, bills
4224	Clerk, hotel front desk
4323	Clerk, air transport operations
4413	Clerk, classification: data processing
5412	Constable
Technical to 1st Language Level	
3135	Caster, central control
Technical to 2nd Language Level	
3131	Dispatcher, load: electrical (power station)
3132	Operator, control-panel: incinerator
3134	Operator, blender: petroleum and natural gas refining
3139	Controller, robot: industrial
3313	Assistant, accounting
3331	Agent, clearing
3333	Agent, employment
3359	Courier, diplomatic
3422	Coach, athletic
3433	Accessioner, library
Technical to 3th Language Level	
1412	Manager, café
1420	Dealer, car: managing and supervising staff
3111	Technician, astronomy
3112	Clerk of works
3113	Estimator, engineering: electrical
3115	Dockmaster, dry: dock
3116	Estimator, engineering: chemical
3117	Acidiser, oil and gas well
3118	Draughtsperson
3119	Investigator, fire
3121	Boss, shift: mining
3122	Coordinator, area: manufacturing
3123	Coordinator, building: construction
3133	Operator, cement production plant
3141	Technician, anatomy
3142	Demonstrator, farm
3143	Technician, arboriculture
3151	Engineer, chief: ship
3152	Captain, port
3153	Astronaut
3154	Controller, air traffic
3155	Engineer, air traffic safety
3213	Assistant, pharmaceutical
3230	Bonesetter
3251	Assistant, dental
3252	Analyst, medical records
3254	Dispenser, optical
3255	Aide, therapist: physiotherapy
3256	Assistant, clinical: helping doctor
3257	Inspector, food sanitation and safety
3258	Ambulanceman
3259	Assistant, speech therapy
3311	Broker, finance
3314	Assistant, actuarial
3315	Adjuster, claims
3324	Broker, commodities
3330	Imputed Category: Special Needs Teachers
3339	Agent, literary
3421	Aerialist, sport
3423	Guide, outdoor adventure
3431	Journalist, photo
3434	Chef
3435	Artist, body
Technical to 5th Language Level	
1411	Director, hotel
3332	Administrator, conference
3334	Agent, estate
3344	Assistant, administrative: doctors surgery
3351	Guard, border
3352	Collector, tax
3353	Inspector, pensions
3355	Agent, inquiry: police
Professional to 2nd Language Level	
2651	Artist, ceramic
2653	Arranger, ballet
Professional to 3th Language Level	
1211	Banker
1311	Manager, agricultural production
1312	Captain, shore: fishing
1321	Director, power station
1330	CIO
1346	Director, bank
2152	Engineer, computer: hardware
2153	Engineer, broadcast
2262	Chemist, dispensing
2354	Coach, vocal
2433	Agent, sales: engineering
2434	Agent, sales: communications (technology)
Professional to 4th Language Level	
110	Admiral
1111	Alderman
1112	Administrator, city
1113	Chief, village
1114	Chairperson, charity
1120	CEO
1213	Coordinator, policy: government
1219	Director, administrative services
1322	Controller, production: mining
1323	Builder, project
1324	Captain, shore: shipping
1341	Director, after school care
1343	Coordinator, community aged care
1345	Academic, university: head of department or faculty
1349	Director, design service
2144	Architect, marine
2145	Consultant, engineering, chemical
2146	Assayer
2151	Designer, engine: electrical
2161	Architect, building
2165	Cartographer
2211	Consultant, medical: general practice
2212	Allergist, clinical
2230	Acupuncturist
2261	Dentist
2263	Adviser, environmental health
2264	Physiotherapist
2267	Optician, ophthalmic
2269	Chiropodist
2411	Accountant
2412	Adviser, debt
2413	Analyst, bond
2652	Accompanist
2656	Anchor, news
2659	Acrobat

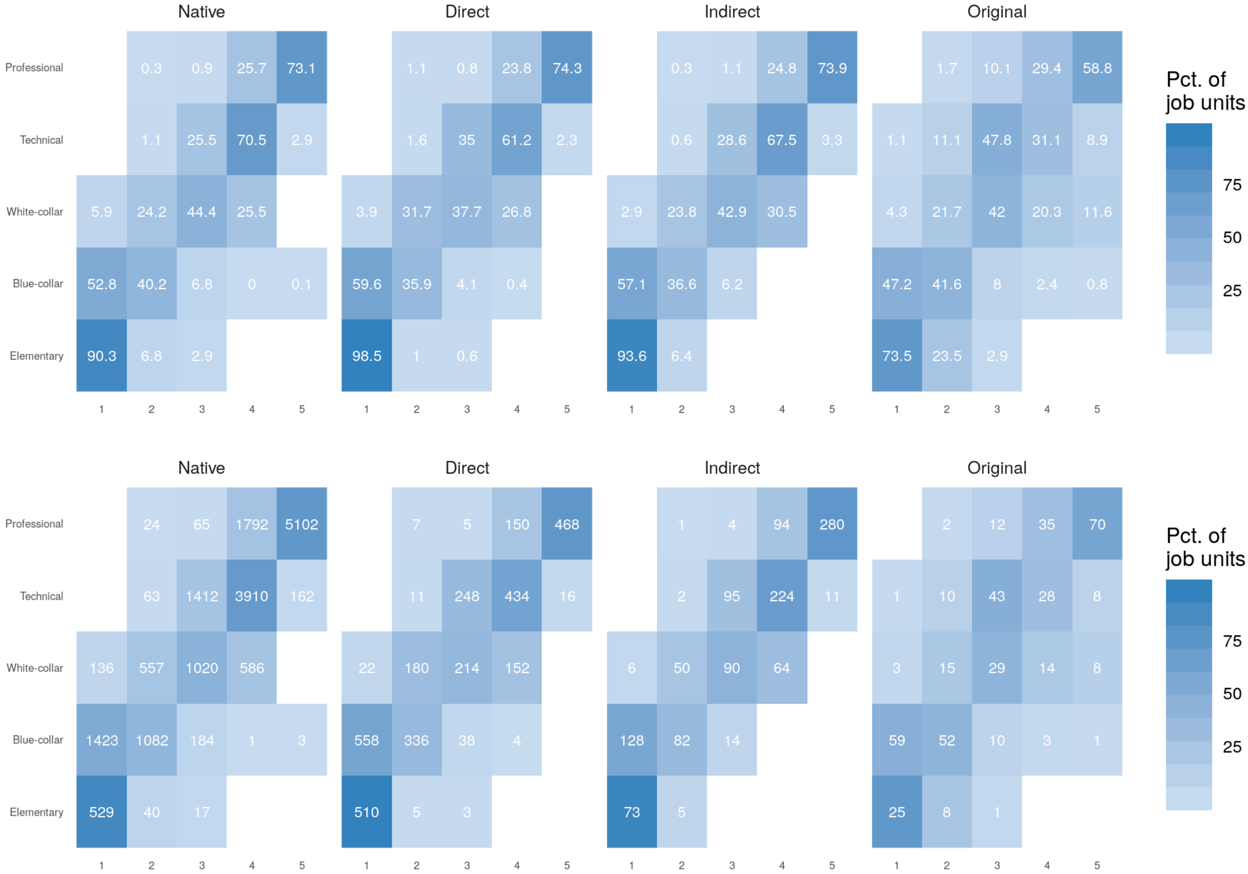


Figure (25) Observed distribution density change induced by the Language Skill reclassification by migration background. Original refers to the number of job units migrated, without sample weightings. (Top) Percentage (Bottom) Number of observations.