

A measurement of migrants' skills across European host countries

Lorenzo Corsini

lorenzo.corsini@unipi.it

Univeristà di Pisa

Irene Martelli

Research Article

Keywords: Migration, item response theory, skills, refugees

Posted Date: September 28th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10907381/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

A measurement of migrants' skills across European host countries

Lorenzo Corsini¹, Università di Pisa, Pisa, Italy

Irene Martelli², Independent Researcher, Pisa, Italy

ABSTRACT

Our paper provides an estimation of migrants' skill and explores its distribution across European host countries. The estimation is based on data from the 2014 ad-hoc module of European labour survey on migrants and it is performed with an item response theory model. We assume that skill is a latent trait which is correlated to several (observable) characteristics. Our results show an ample variation of migrants' skills and we find that most of Nordic countries, Switzerland and UK attract and host the best qualified pool of migrants. From a policy point of view, we acknowledge that the large differences in the distributions of skills are difficult to ignore when aiming at a coordination of European immigration policies. On the other side, these ample differences can be analysed to understand which policies were more successful in attracting and withholding high-quality migrants.

Keywords: Migration; item response theory; skills; refugees.

JEL codes: F22; J24; C40

Corresponding author: Lorenzo Corsini, lorenzo.corsini@unipi.it

¹lorenzo.corsini@ec.unipi.it

² irene.martelli@gmail.com

Clinical trial number: not applicable

Ethics Declaration

This study was conducted in accordance with the principles of research integrity set out in the European Code of Conduct for Research Integrity. The analysis relied exclusively on anonymized secondary data, and the authors had no access to directly identifiable personal information. Therefore, ethics committee approval and additional informed consent were not required.

Funding

The authors received no funding for this research.

Competing interests

The authors declare that they have no competing interests.

1. Introduction

One of the dimensions of the migration process that appears to be important in understanding the causes and the impact of this phenomenon is related to the composition of migrants in terms of skills, abilities and expertise. It is obvious that, depending on the skills of people that moves to a given country, the impact on the host economy may differ greatly: highly skilled migrants may boost efficiency and innovation, whereas low skilled migrants may be perceived as a burden. This latter negative view can however be reversed in the presence of shortage of labour supply in low skill occupations. On similar notes, migrants with high skills may generate high income, greatly contributing to the fiscal and welfare systems, whereas low skilled may end up being a burden in terms of public budget (see Chiswick, 2005). Given the importance of migrants' skills for host economies, it becomes crucial to understand how these skills are distributed across different countries and what factors may drive such differences. This very reasoning is often present in the current policy debate, both because it defines the overall costs and benefits of hosting migrants and because it rises concerns on which countries gain or lose from the migration process.

On one side, it is evident the gain from the high skill migration, so much that, within the European Union context, the European Council issued the Council Directive 2009/50/EC that states that "it is necessary to facilitate the admission of highly qualified workers and their families by establishing a fast-track admission procedure and by granting them equal social and economic rights as nationals of the host Member State" and, even more recently, produced a proposal to "substantially broaden its [of Directive 2009/50/EC] impact in attracting more highly skilled workers to the EU" (European Commission, 2016). On the other side, some national politicians complain that their own country suffer from the bad selection, attracting only "bad" workers, whereas "good" workers moves on to other countries.

The issue of selection and self-selection of migrants is not merely a political discussion: it was introduced in economic terms in a popular paper by Borjas (1987) where a model is used to show

how migrants may self-select on the base of their skills, with better migrants choosing countries with more inequality (where they can potentially obtain better incomes) and with worse migrants choosing countries with less inequality (where they are somehow guaranteed to obtain a relatively high income). The same reasoning applies to destination countries with more generous welfare systems, which should attract “worse” migrants.

Our contribution starts from these considerations and uses detailed information from a survey on foreign born citizens to produce an estimation of a (latent) variable representing the skill of an individual (migrant in our case). Once we have obtained this estimation, we examine how this skill variable is distributed among the European host countries.

Our idea is that the concept of skills is an unobservable/latent trait that is correlated to other observable characteristics (education, occupation, language proficiency just to name a few) and, accordingly, we will adopt an item response theory (IRT) model to estimate this latent trait. This method use qualitative information gathered in “items” that usually comes from surveys or tests, to obtain an estimation of a latent variable that is related to the observed items: in our specific case we use data from the 2014 EU-Labour force survey which contains an ad-hoc model on “the labour market situation of migrants and their immediate descendants”. This dataset allows us to identify migrants, country of origin, their refugee status as well as several other relevant characteristics that we believe to be correlate to their “skill”.

Once we have obtained an estimation of this skill variable, there are a number of aspects that we want to specifically address.

- 1) We want identify which European host countries have migrants with high (and low) average values for the skill variable.
- 2) We explore whether the observed cross-country distribution of migrant skills is descriptively aligned with theoretical selection mechanisms discussed in the literature. To do that we compare the

average skills of migrants with income, income inequality and social welfare spending in a given host country.

3) We want to focus on specific groups of migrants: individuals moving from (extra-EU) poor countries and individuals that moved seeking political asylum. We aim to assess how these groups of migrants fare in terms of skills compared to the rest of migrants and how their average skill level is distributed across host countries.

All these aspects are extremely important from a policy point of view. The actual distribution of migrants' skills across host countries is something that is hard to ignore when trying to coordinate European migration policies. Moreover, differences in this distribution can be exploited to assess which migration policies were successful in attracting good quality migrants. Clearly, we are not stating that the average skill level of migrants depends solely on migration policies: rather, we are implying that it is a combination of several factors (economic attractiveness, geography, culture) and that the migration policy is one of them

The analysis of migrants' skills and their distribution has received attention within the past literature even if usually the skill dimension was solely represented by educational attainment.

The already mentioned paper by Borjas (1987) set the discussion on the self-selection phenomenon. Some later papers appear to confirm, at least theoretically, the self-selection (see Katz and Stark, 1987 and Chiswick, 1999).

A relevant analysis of immigrants' skills across Europe is contained in Dustmann and Frattini (2013): in their contribution the skill dimension is represented by the educational attainment, occupation and income and these aspects are separately examined drawing comparison between native and migrants. Their findings show how the Scandinavian countries, Ireland and UK attract highly educated migrants (something that was also confirmed in Dustmann and Glitz, 2011). In an attempt to push forward these findings, in our work we will try to estimate a single variable capturing the skill dimension: the

use of the 2014 EU-LFS module on immigrants allows us to have several information and variables that can be exploited to obtain a more precise measure for skills.

Another ample analysis of migrants' skills, and high skill in particular, is contained in Huber et al. (2010): even if their work use mostly education to identify skills, they also try to explore the over-qualification phenomenon among native and migrants and find that the latter are more likely to end up in an occupation for which they are over-qualified. A rare work that tries to expand the definition of skills beyond education attainment is Dustmann and Fabbri (2003): their contribution analyses the role of language proficiency of migrants in the UK and finds that better proficiency has a positive impact on employment outcome and earnings, indicating thus that this specific ability is actually important to obtain a realistic assessment of skill. The relevance of language in determining labour market outcomes is also highlighted in Drinkwater (2017).

Some works on the distribution of skills within the country of origin (whereas we want to compare the skills of migrants across destination countries). Among them, an analysis of skills (represented by educational attainment) is present in Peri (2016): in this contribution the distribution of skills within the country of origin is examined and the findings show that a positive selection is at work so that the most educated are more likely to migrate. In a similar perspective, the work by Hatton and Williamson (2006) defines skills on the base of occupation and schooling and looks at their distribution within the origin countries and find that, in the long run, there has been reduction of the quality of individuals that choose to migrate, so that in recent decades the migrants are now less qualified than they used to be in the past.

If we consider the bulk of work represented by the above literature, we can see that in most of the cases the definition of migrants' skill directly coincides with educational attainment: our use of the LFS ad-hoc module allows us to observe many more variables related to skills and we use them to estimate a single latent variable measuring skill: we believe that this provide a valuable contribution on the topic of migrants' skills.

Apart the direct analyses of migrants' skills, there is a strand of literature that tries to investigate the relationships between immigration policies and the selection in terms of skills. Among them, Borjas (1993) compares immigrants in the US and Canada and finds that the latter attracted more educated immigrants after the introduction of a point system that takes into account characteristics (and skills) of the applicants for visa. Antecol et al. (2003) analyse immigration policy Australia, Canada and US: their findings show that Australia and Canada use an immigration policy that takes into consideration skills when granting admission and this translates in their immigrant having better English fluency, education and income when compared to US. However, they discuss that this may also be due to the higher presence of Latin workers that moves to the US simply for vicinity reasons. The work by Cobb-Clark (2004) focuses on the Australian case and finds that changing the admission procedure toward a system that is based on skills produced a pool of immigrants that is more educated, more fluent in English and that has more pre-migration labour market experience: these improvements in human capital endowments (combined with the adoption of income-support policy) translated in better labour outcomes both in the short and in the medium run.

Within the analysis of the migration process, the issue of refugee migration is becoming particularly relevant given the recent dramatic world events that forced a surge in this kind of migration (see Hatton, 2016). In terms of the characteristics and economic outcomes of refugees, Cortes (2004) looks at refugees waves from the eighties and finds that refugee immigrants works and earns less than economic immigrant though a catch-up process appear to be a work. A similar catch-up process was found, for Sweden, in Luik et al. (2016) and in Europe as a whole in Aiyar et al. (2016).

More recently, the situation of refugee migrants and their labour market characteristics is assessed in an ample contribution by Dustmann et al. (2017). In particular, they show that past waves of refugees displayed lower educational level than extra-EU immigrant. They also confirm this fact using more recent data on refugees that moved to Germany from Afghanistan, Iraq and Syria and that display low education. Moreover, they find that, even after controlling for this worse education level, refugees

experience a slower labour market integration relative to non-refugee migrants. Our contribution wants to provide a more precise assessment of refugee skills, producing an actual measurement and analysing their distribution across countries of destination.

The rest of work is organized as follows: section two presents our estimation methodology and describes the data we use; section three produces the main estimations and describes the main results in terms of average mean skill and its distribution across host country; section four focuses on two specific groups of migrants, those from non-western extra-EU countries and those that are refugees; section 5 discusses our findings and concludes.

2. Statistical methodology and data description

The aim of our paper is to estimate a variable that represents the skill level of migrants and, to this end, we construct a comprehensive measure of migrants' skills that allows for meaningful comparisons across host countries. In this perspective we refer to the term "skill" as something that signals the "ability and capacity to carry out processes and to be able to use one's knowledge in a responsible way to achieve a goal" and that can be declined in three different types: "cognitive and metacognitive; social and emotional; and practical and physical" (OECD, 2019). Clearly, this concept is not directly measurable and to all extents we consider skill as an unobservable and latent variable reflecting multidimensional components. For these very reasons we adopt item response theory, a statistical technique that is able to measure latent characteristics. In practice, we assume that the skill, even if latent, is associated to some observable "items" that are thus correlated to it. These items are measured through surveys (or tests in other cases) and their values are strictly correlated with the skill. In our specific case, the observed items represent the characteristics of individuals in terms of their education, their background and the capacity that they show in their occupation: we assume that all these characteristics are correlated with a latent trait that we call "skill". Some of these

characteristics are measurable by binary variables (having or not having a certain attribute) while others represent grades reached in a certain characteristic, so that higher values represent better grades and should thus be associated to higher skill. In what follows we introduce the IRT methodology and we then describe the data and how we use them within our IRT model estimation.

2.1 Item response theory

Item response theory was developed as a way to measure latent constructs. The concept of latent construct is related to something that can only be assessed indirectly, using observable variables that are somehow related to the construct: typically, a test or a survey that provide some observed variables (called items) is given to a sample of individuals to assess a unobservable characteristic (that is, the latent trait).

IRT models were initially introduced to deal with educational and psychometric assessment issues (Lord and Novick, 1968; Minton et al, 1988). In facts, in these area, the latent aspect of the object of investigation (psychological profile and individuals' ability) and the test characteristics are relevant at the same time. In such a context, IRT represents the main method to obtain an estimation of both the individuals' score and the item properties (see Hambleton and Swaminathan, 1985; van der Linden and Hambleton, 1997). This methodology is often used to measure abilities in cognitive contexts (see Schmidt. and Embretson 2003), but it can also be used in social science to provide an assessment, for example, of quality of life (Cella et al., 2002), deprivation (Szeles and Fusco, 2013), tourism inclination (Martelli et al., 2016) and financial literacy (Walstad, 2017). To our knowledge this is the first time that it is applied in the analysis of human capital and work skills.

The main advantages of IRT consists in the simultaneous assessment of latent trait and of the relevance of the items. Given the data available and the aim of our study, we use the graded response model (GRM; Samejima, 1969) which is an extension of the two-parameters logistic model (2PLM; Birnbaum, 1968).

Let us consider a random vector, $\mathbf{Y}_i$, of p item responses for the i -th individual ($i = 1, \dots, n$) and the observed responses, $\mathbf{y}_i = (y_{i1}, \dots, y_{ip})$. Let's denote θ_i as the latent trait of the i -th individual, which is assumed to follow a standard normal distribution..

2.1.1 The two-parameter logistic model

The 2PLM is used for binary items, e.g. correct/incorrect, and specifies the probability π_{ij} of a correct response from the individual i to the item j , as a logit function of the predictor η_{ij} ,

$$\pi_{ij} = P(Y_{ij} = 1 | \theta_i) = \frac{\exp(\eta_{ij})}{1 + \exp(\eta_{ij})} \Rightarrow \text{logit}(\pi_{ij}) = \eta_{ij} \quad (1)$$

which depends on θ_i and on the vector $\boldsymbol{\xi}_j = \{\alpha_j, \beta_j\}$ of parameters that characterize the item j ($j = 1 \dots p$). Each item j is defined by two kinds of parameters: the *difficulty parameter* β_j and the *discrimination parameter* α_j . The predictor (1) can be written as $\eta_{ij} = \alpha_j(\theta_i - \beta_j)$ and model (1) becomes the 2PLM (2):

$$\text{logit}(\pi_{ij}) = \alpha_j(\theta_i - \beta_j) \quad \theta_i \sim N(0,1). \quad (2)$$

2.1.2 The graded response model

The GRM for ordinal items (Samejima, 1969) represents a generalization for ordinal data of the 2PLM. Graded responses can be seen when we observe Lykert-type scales (for example, from “fully disagree” to “fully agree”) and response that have a order according a range of scores.

There is a neat difference in the way IRT models for polytomous items operate with respect to the binary models: basically, in the presence of ordered polytomous items we can treat them as “concatenated dichotomous” items, so that we can actually implement a dichotomizations of item response data.

Let consider a set of p ordinal items, Y_j , each with K_j ordered categories, indexed by k . Therefore, we can describe each item by $K_j - 1$ boundaries or thresholds: $\kappa_{j,1}, \dots, \kappa_{j,K_j-1}$. The probability to reach a given category k or higher, it increases monotonically with the growth of the latent trait and the thresholds satisfy the order constraint: $\kappa_{j,1} < \dots < \kappa_{j,K_j-1}$.

As for the implementation of the dichotomization, the GRM has a specification that takes into account the probability that the answer will fall in *category k or higher*: the probability π_{ijk} that the i -th individual will be part of the k -th category on item j is calculated as the difference between the probability of responding higher than the lower threshold for the category $\pi_{ij,k}^*$ and the probability of responding higher the category's lower boundary $\pi_{ij,k+1}^*$:

$$\pi_{ij,k} = P(Y_{ij} = k | \theta_i) = \pi_{ij,k}^* - \pi_{ij,k+1}^* \quad \theta_i \sim N(0,1) \quad (3)$$

where $\pi_{ij,1}^* = 1$, $\pi_{ij,K_j+1}^* = 0$ and $\pi_{ij,k}^* = P(Y_{ij} \geq k | \theta_i) = \frac{\exp(\alpha_j \theta_i - \kappa_{j,k})}{1 + \exp(\alpha_j \theta_i - \kappa_{j,k})}$ (Samejima, 1969, 1996).

The discrimination parameter α_j determines the response functions' slope and does not change between all the responses' category of a given item. This property guarantees that non-negative probabilities exist (Steinberg and Thissen, 1995). The boundary parameters $\kappa_{j,k}$ change within a given item, according to the order constraint $\kappa_{j,k-1} < \kappa_{j,k} < \kappa_{j,k+1}$, and at each level of $\theta = \kappa_{j,k}$ the individual has a probability of 0.5 of fall in the category (Reeve, 2002). For binary items, the specification of GRM coincides with the 2PLM.

The methodological framework described in the previous paragraph fits both the scope of our study and available data type (binary and ordinal).

2.2 Data description

Our analysis relies on the data contained in the 2014 ad-hoc module of the European Labour Force Survey (LFS): this specific module was focused on “the labour market situation of migrants and their

immediate descendants”. This module allows us to focus on information that are particularly important in assessing the actual skills of migrants and that are solely contained in the ad-hoc module. In the rest of the work, we will refer to migrants as the individuals that were born in a different country from the one of residence.

The data we use cover several European countries: Austria, Belgium, Bulgaria, Czech Republic, Finland, France, Greece, Hungary, Italy, Norway, Poland, Portugal, Romania, Spain, Sweden, Switzerland and United Kingdom. Some countries are not part of this group because they are not included in the ad-hoc module (Denmark, Germany, Ireland and the Netherlands) while we excluded the rest of European countries because the sample size of migrants, in those countries, was negligible. In total, we have 32256 observations.

It is important to stress that the estimation of the latent skill variable for migrants was obtained using observations from all the above selected countries. However, when we report detailed statistics for migrants or sub-groups of migrants in a given host country, we will restrain from reporting results when the sample size in that country is too small.

In order to estimate the IRT model, we need to use variables that are related to skills: these variables represent the “items” that are correlated with the very skills. In other words, our assumption is that there exists a latent trait (skill in our case) that affect all the items that we are going to introduce. In our specific case, some of these items are available for all the migrants whereas some items may be available only for sub-groups of migrants. The main sub-group is represented by the employed: in fact, items describing the current occupation (and thus related to skills) are only available for employed migrants. Luckily, the IRT model, can overcome this aspect because when an observation

has an item that is not available³, it excludes it from the likelihood calculation but keep using the non-missing items for that observation (see Schmidt and Embretson 2003).

The list of items that we assume to be affected by the latent trait (skill) and that we will use in its estimation are reported in Table 1: we derived the items from an elaboration of the information contained in the ad-hoc module of LFS. We coded all the items so that higher values are associated to what should reasonably signal better skill. The only exception to this rule is gender: in this case, we tried two different coding, one in which being female is associated to better skill and the other in which we try the opposite.

TABLE 1

These are all the items that we assume to be associated to the latent trait “skill”: however, this does not mean that they are actually associated to skill nor that our coding truly reflects the real skill. For these reasons, in the next section, we will test if all the above items are indeed correctly related to the latent trait and we will exclude the items that do not pass the test. The items that pass this selection, will be used to estimate the latent trait.

3. Estimation of the skill latent variable

We proceed now to estimate the skill latent trait using the IRT model: these estimates allow us to highlight substantial differences in migrants’ skill composition across European countries. In principle, we could use all the items that we previously described: however, we need to check for the appropriateness of the items and to do that we apply the Mokken Scale Procedure (MSP): this procedure allows to select/exclude the items on the base of an appropriateness scale measured by

³ To better understand this we can draw a comparison with the assessment of the score of a school test: if in our case one item is missing for an observation this can be seen as a question that, in a school test, was never asked to that individuals. Clearly, the non-asked question does not affect the final score in the school test.

Loevinger's H coefficient (see Hardouin et al., 2011). We also compute Cronbach's alpha to assess internal consistency between items (see Cronbach, 1951). The results from the implementation of MSP to our dataset resulted in the exclusion of "Gender" (in both its specifications), "Current job is not on fix term base (i.e. it is open ended)" and "Over-qualified for current job". Basically, this means that these items are not related to the others and that they appear not to be affected by the latent trait. Our final set of items has good scale properties (Loevinger's H coefficient is 0.5) and reasonable internal consistency (Cronbach's alpha is 0.56).

We now proceed in the estimation of the skill latent variable and of the discrimination and difficulties parameters associated to each item: all estimations were obtained using sample weights provided by the LFS. In terms of interpretation, the discrimination parameters tell us how strong the effect of skill is on this specific aspect so that, whenever the skill changes, the item response is very likely to change too. As for the difficulty parameters, they signal how difficult is, even for individual with high skill, to obtain a high value for these items. All the parameters are reported in Table 2.

TABLE 2

Our estimations point to the fact that migrants' skill is particularly affected by education, parent education and by the best occupation ever reached. To a lesser extent skill is associated to the supervisory responsibility and to the technical and scientific fields of study. As for the difficulty parameters, we observe a very high value for the technical and scientific fields of study: this means that only individuals with very high skill usually have these characteristics. In other words, when we observe individuals whose field of study falls in that category, we can be fairly certain that their skill level is high, however the reverse is not necessarily true: not having this characteristic does not necessarily mean that the individuals have low skill level. Interestingly, language proficiency has low difficulties values (especially for the first parameter): this means that is quite easy to reach some knowledge in the language, but this is not necessarily a sign of skill.

The estimation allows also a direct estimation of the skill latent trait: the estimation normalize to 0 the mean value of this trait across the whole population but we can report the mean value of the skill trait in each host country. We depict these values in Figure 1 where we report only countries where the sample size of migrants is at least 150.

FIGURE 1

The figure depicts a quite clear situation with an evident North-South gap in terms of average skill level of immigrants; moreover, eastern countries perform well. In particular migrants in Switzerland, UK, Norway and Czech Republic have the largest endowment in terms of skills.

We can compare our results with a benchmark represented by schooling attainments of migrants in host countries as reported in Dustmann and Frattini (2013). In accordance with their findings we still find that Norway, Sweden and UK host migrants whose average skill is high. On the contrary, in Finland, we find a lower skill level than the direct school attainment would have predicted. On similar notes, migrants in Spain are doing worse than expected and in Portugal are doing better. The disappointing evidence for Italy is partly confirmed, though we probably find an even more grim result than the one obtained considering sole education. In addition to that benchmark, we find very large skill endowments in Switzerland and we also observe that the skill level of immigrants in Czech Republic and Hungary is high (eastern countries were not reported in Dustmann and Frattini, 2013): these latter results should be read carefully because the number of observation in these countries is not large (in Hungary in particular) and because this might be due to the composition of migrants in terms of country of origin. We will better examine this issue in later sections.

Apart the visualization of the mean values of skill level, we want to statistically test if these estimated values are different from country to country. The estimated mean values, their standard errors and the results of the pairwise Wald for two countries at a time are reported in Table 3.

TABLE 3

The results of the tests are not surprising: in most cases the differences in the mean values are statistically significant across countries. However, there is a group of high skill countries (Czech Republic, Hungary, Norway, Sweden and UK) whose mean values are similar and not statistically different one from the other.

3.1. Average skill level and the selection of migrants

We want now to provide some elements that could be correlated with the observed distribution of skills across host countries. Bearing in mind the possible selection mechanisms of migrants, we provide a comparison of average skill level with 4 elements in each country: i) the average GDP per capita (in PPP), ii) income distribution in terms of Gini index, iii) income distribution in terms of the 80/20 income ratio and iv) social spending as a share of GDP. The results of this comparison are contained in Figure 2.

FIGURE 2

Starting with the relationship between migrants' skill and GDP per capita (upper left panel), we observe what it looks like an increasing trend. The pattern is even more clean-cut if we consider the eastern countries as possible outliers. The message is clear: richer countries host the best migrants.

The relationship between skill and inequality measures (upper right and bottom left panels) is also quite evident: it looks like a negative one. That is, contrary to standard selection pattern, worse migrants do not go (or are not hosted) in countries with smaller inequalities. In truth, this is partly due to the fact that, in Europe, richest countries are also displaying lower inequalities, so that the selection mechanisms based on inequality may be completely overshadowed by a mechanism that brings best workers towards richest countries. Moreover, it is important to bear in mind that we are looking at a cross-section of data and the pattern may be different in an evolutionary perspective.

Finally, the relationship between skill level and social expending is not clear: there may be a vague decreasing pattern, but it certainly disappears if we consider eastern countries as outliers.

4. Results for extra-EU and refugee migrants

The previous section dealt with the cross country distribution of skills of all migrants and shows an ample variation from country to country. However, this large variation may be due, in part or fully, from the fact that some host countries are more likely to attract migrants from specific country of origin and this origin may explain a large part of the skill level of migrants. To assess this possibility and to produce a more homogenous comparison we now restrict the analysis to specific groups of migrants: i) the non-western extra-EU individuals and ii) migrants that move to seek asylum and protection.

The first group represents probably one of the most problematic in terms of integration: they usually come from disadvantaged or poor countries and they are likely to meet severe difficulties in their working life. The focus on the second group is also particularly relevant: the current world situation has brought again a large surge of refugees and those that move for these reasons come, by definition, from countries facing unrest, wars and severe violations of human rights so that they are likely to have a very difficult past.

4.1 Non-western extra-EU migrants

Individuals from these countries often come from developing or transition countries. We define this group excluding migrants from EU, from North America and Oceania: the bulk of them comes from Africa, Asia and from the less developed areas of Europe but also from South America. The average value of skill for this group of migrants is -0.14 which is statistically different from the skill of the rest of migrants from EU and other Western countries (whose mean is 0.201): full data on this comparison is available in appendix A. This result is not surprising, however, what we want to

examine is whether this differential is also observed within each host country. To do this we report in Figure 3 the average skill in each host country.

FIGURE 3

The distribution of skill level across host countries more or less reflects the distribution that we observed for the full sample of migrants. There is, however, a striking element: in UK, this group has the same average skill level that was observed for the whole sample; apparently, in this country, even migrants from disadvantaged countries appear to be well selected. A few more differences exist: we observe a large drop in skills for this group in Switzerland and similar evidence is found in Austria; on the contrary, in France and Portugal the drop is relatively small. Finally, it is interesting to note that the good quality of migrants in Czech Republic is confirmed also when we restrict the analysis to this group⁴.

Combining this evidence with previous results from Figure 3 we can provide an indication that, apparently, Switzerland is particularly attractive and selective in terms of country of origin whereas UK is more attractive and selective in terms of actual skills rather than in the origin⁵.

We provide in Table 4 the detailed estimated mean values, their standard errors and the results of the pairwise Wald test for two countries at time.

TABLE 4

The tests perform as expected: the only noteworthy result is that the drop in skills in Switzerland is confirmed also by the fact that the values of average skill level are statistically different with respect to the other countries (that is, skills in Switzerland are now lower than in some other countries).

⁴ We have not reported the value for Hungary because the sample size was too small: however, the fall in migrants' skills for this group was particularly severe (from 0.3 to 0.08). In any case, the sample size suggests applying caution in considering this result.

⁵ In principle, the result of UK could also be explained by the large presence of Indian migrants that could be associated with good skills. However, if we exclude the Indian migrants from the estimation for UK, the mean value of skill level for non-western extra-EU migrants remains almost the same.

4.1 Refugee Migrants

The last group that we want to focus on is comprised by the migrants that declared that the main reason for their move was to seek “International protection or asylum”. Focusing on this groups causes a reduction of the sample size: the overall sample is still reasonable (928 observations across all the countries) but it obviously shrinks even more within each single country.

This said, we the estimated average skill level for refugee was 0.131 whereas the average level for non-refugee is 0.004 and the difference is statistically significant (full data is available in appendix A). This result, however, could be driven by the origin of refugee migrants that often come from disadvantaged countries rather than from being refugees. For this reason, to produce a more homogeneous comparison, we restrict the analysis only to migrants from non-western extra-EU migrants and compute their average skill level distinguishing between non-refugee and refugee: in this case we find that skill level for refugees is -0.259 while for non-refugees is -0.137 and these values are statistically different (full data is available in appendix A).

Even in this case the refugees display a strictly lower average skill value. This signals the extreme disadvantage of this group and calls for attention when designing policies to host and integrate them.

Finally, we report data on skill level of refugees within each host country. As mentioned above, the sample size now becomes small and we decided to be less strict in the minimum requirement in terms of sample size, choosing 40 observations as the threshold: clearly, great caution should be applied when reading results for country whose sample size is particularly small. The adoption of this thresholds reduced in any case the number of countries that we can examine, bringing their number to 9. Results are reported in the Figure 4.

FIGURE 4

The results appear quite surprising, often reshuffling evidence found when examining the full sample of migrants. The case of Switzerland is particularly striking: the refugees in this country has an extremely high average skill level and this is even more surprising given the low skills that migrants from disadvantaged country had (see Table A in appendix A). The truth here is that Switzerland is quite an outlier in terms of refugees: in our sample, more than half of them comes from the EU; no other host country displays this specific characteristic. It is thus likely that the attractiveness of Switzerland and its asylum policy generate an extremely positive selection process. This is partly true also for France that displays a pool of refugees that has quite high skills compared to the rest of the countries. The reverse is true for UK and Finland: these nations had, respectively, very high and average skill level among the whole group of migrants from disadvantaged countries but they instead have low skills among the refugees. Finally, Italy is still in the bottom positions of the distribution even if, considering migrants from disadvantaged countries, the refugees and non-refugees display similar skill levels.

In any case, the above results were obtained from small samples and to provide information of the statistical significance of the differences across countries we report in Table 5 the detailed estimated mean values, their standard errors and the results of the pairwise Wald tests for two countries at time.

TABLE 5

The results show that, when considering pairwise comparisons between countries, differences are not always significant. What can be safely asserted is that there appears to be roughly three groups of countries: there is a group (Switzerland and Norway) where refugees have generally higher skills than in the rest of countries; then there is a group (Belgium, Italy and Finland) where refugees' skills are particularly low and, finally, there is a group of countries that falls somewhere between the others (France, Sweden, Austria and UK).

5. Discussion and conclusions

Our analysis provides a measure of migrants' skills looking at their distribution across European host countries. The measure of skill that we estimated stem from a multidimensional concept of capabilities and abilities and it is seen as a latent trait that is associated to several components.

Our findings show that the average skill of migrants displays ample variation across host countries, with some countries able to attract highly skilled migrants (most of the Nordic countries, Switzerland and UK) whereas others host mostly low skilled (the southern countries and also France). We compared this ample variation with some macroeconomics variables (in term of GDP, income distribution and social spending) of host countries and we noticed that the pattern appears to follow the fact that richer countries hosts the high quality workers. In this sense, the role of income distribution is overshadowed: in fact, in Europe, countries with high GDP per capital usually have less inequality so that we concurrently observe that best migrants reside in countries that are richer and less unequal. From this point of view, social spending does not play a clean-cut role.

The large differences across countries provide some insights from a policy point of view. If a coordinated European policy on immigration has to be discussed and conceived, we have to acknowledge that some countries probably gains more than others from migration processes. Moreover, our data show that southern countries attract and withhold low quality migrants: if this geographical area is the main port of entry of migrants, this clearly constitutes a problem for them but also for the whole European countries. In fact, the countries that act as ports of entry may fail to attract good migrants given that the first step takes place in what appear to be unattractive countries. This is particularly worrisome given the European Commission declared aim of attracting high quality migrant. This is particularly worrisome given the European Commission declared aim of attracting high quality migrants. At the same time, these differences suggest that national contexts play an important role in shaping the composition of migrant populations. In particular, variations in migrants' skills across countries may reflect not only economic conditions but also institutional

settings, migration policies, and labour market structures. Understanding these differences is crucial for designing effective and coordinated migration policies in Europe. In this context, an analysis of the characteristics and of the policies adopted in countries that are successful in selecting good migrants may provide further information on how to design adequate migration policies.

Our work also provides a deeper analysis on migrants from disadvantaged countries and on refugees. The analysis of the former in part confirm the first part of our analysis. However, a few elements change: the most striking case is UK that appears to produce an extremely good selection within this disadvantaged group.

The results on the political refugees instead produce novel insights, even if the small sample size calls for caution in interpreting our results. What we can safely state is that we confirm previous evidence on the fact that refugees have lower skills than the rest of migrants. However, we also find that their cross country distribution does not mirror the distribution observed for the rest of migrants. From this point of view, we believe it would be useful to develop further research in two directions. On one side, a larger dataset on refugees would allow to obtain more precise measures and to pinpoint exact situations. On the other, a comparing the skills and quality of refugees with country policies and the requirement to grant the asylum status would shed light on which policies appear to be effective.

REFERENCES

- Aiyar, S., Barkbu, B., Batini, N., Berger, H., Detragiache, E., Dizioli, A., Ebeke, C., Lin, H., Kaltani, L., Sosa, S. and Spilimbergo, A. (2016). The refugee surge in Europe: Economic challenges. *National Institute Economic Review*, 235(1), F16-F31.
- Antecol, H., Cobb-Clark, D.A., and Trejo, S.J. (2003). Immigration policy and the skills of immigrants to Australia, Canada, and the United States. *Journal of Human Resources*, 38(1), 192-218. doi:10.2307/1558761.
- Birnbaum, A. (1968), Some latent trait models and their use in inferring an examinee's ability. In *Statistical theories of mental test scores*, by Lord, F.M. and Novick, M.R., Addison-Wesley.
- Borjas, George J. Self-Selection and the Earnings of Immigrants. *The American Economic Review* 77, no. 4 (1987): 531–53.
- Borjas, G.J. (1993). Immigration Policy, National Origin, and Immigrant Skills: A Comparison of Canada and the United States. In *Small Differences That Matter: Labor Markets and Income Maintenance in Canada and the United States*, by Card, D. and Freeman, R., National Bureau of Economic Research, Inc. 21-44.
- Cella, D., Chang, C.H. and Heinemann, A.W. (2002). Item Response Theory (IRT): Applications in Quality of Life Measurement, Analysis and Interpretation. In *Statistical Methods for Quality of Life Studies*, by Mesbah, M., Cole, B.F. and Lee, M.L.T., Springer, Boston, MA.
- Cobb-Clark, D.A. (2004). Selection Policy and the Labour Market Outcomes of New Immigrants. IZA Discussion Papers 1380, Institute for the Study of Labor (IZA).
- Cortes, K.E. (2004). Are refugees different from economic immigrants? Some empirical evidence on the heterogeneity of immigrant groups in the United States. *The Review of Economics and Statistics*, 86, 465–80.

Council of the European Union (2009). Council Directive 2009/50/EC of 25 May 2009 on the conditions of entry and residence of third-country nationals for the purposes of highly qualified employment, Official Journal of the European Union, L 155, 17–29.

Cronbach, L.J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16, 297-334. doi: 10.1007/BF02310555

Drinkwater, S. (2017). Why does unemployment differ for immigrants?. *IZA World of Labor*, 376. doi: 10.15185/izawol.376

Dustmann, C. and Frattini, T. (2013). Immigration: The European Experience. In *Immigration, Poverty, and Socioeconomic Inequality*, by Card, D. and Raphael S., Russell Sage Foundation. 423-456.

Dustmann, C., Fasani, F., Frattini, T., Minale, L. and Schönberg, U. (2017). On the economics and politics of refugee migration. *Economic policy*, 32(91), 497-550.

Dustmann, C. and Glitz, A. (2011). Migration and education. In *Handbook of the Economics of Education*, Elsevier. Vol. 4. 327-439.

European Commission (2016). Proposal for a directive of the European Parliament and of the Council on the conditions of entry and residence of third-country nationals for the purposes of highly skilled employment, COM/2016/0378 final - 2016/0176 (COD).

Hambleton, R.K., and H. Swaminathan (1985), *Item Response Theory: Principles and Applications*. Springer Science & Business Media.

Hatton, T.J. (2016). Refugees, asylum seekers, and policy in OECD countries. *American Economic Review*, 106(5), 441-45.

- Hatton T.J. and Williamson, J.G. (2006). International Migration in the Long Run: Positive Selection, Negative Selection, and Policy. In *Labor Mobility and the World Economy*, by Langhammer, R.J. and Foders, F., Springer, Berlin, Heidelberg. 1-31.
- Huber, P., Landesmann, M., Robinson, C. and Stehrer, R. (2010). Migrants' Skills and Productivity: A European Perspective. *National Institute Economic Review*, 213(1), R20-R34.
- Katz, E. and Stark, O. (1987). International migration under asymmetric information. *The Economic Journal*, 97(387), 718-726.
- Lord, F.M. and Novick, M.R. (1968). *Statistical theories of mental test scores*. Addison-Wesley.
- Martelli, I., Matteucci, M. and Mignani, S. (2016). Bayesian estimation of a multidimensional additive graded response model for correlated traits. *Communications in Statistics-Simulation and Computation*, 45(5), 1636-1654.
- Minton, H.L. (1988). *Lewis M. Terman: Pioneer in Psychological Testing*. New York University Press.
- OECD (2019). *OECD future of education and skills 2030: OECD learning compass 2030*.
- Peri, G. (2016) Immigrants, Productivity, and Labor Markets. *Journal of Economic Perspectives*, 30(4), 3–30.
- Reeve, B.B. (2002). *An introduction to modern measurement theory*. National Cancer Institute.
- Samejima, F. (1969). Estimation of latent trait ability using a response pattern of graded scores. *Psychometrika Monograph Supplement*, 17.
- Samejima, F. (1996). Evaluation of mathematical models for ordered polychotomous responses. *Behaviormetrika*, 23(1), 17-35.

Schmidt, K. M. and Embretson, S. E. (2003). Item Response Theory and Measuring Abilities. In Handbook of Psychology, by Weiner, I. B. doi:10.1002/0471264385.wei0217

Szeles, M.R. and Fusco, A. (2013). Item response theory and the measurement of deprivation: evidence from Luxembourg data. *Quality & Quantity*, 47(3), 1545-1560. doi: 10.1007/s11135-011-9607-x.

van der Linden, W.J. and Hambleton, R.K. (1997). Handbook of modern item response theory. Springer Science & Business Media.

Walstad, W.B. and Rebeck, K. (2017). The Test of financial literacy: development and measurement characteristics. *The Journal of Economic Education*, 48(2), 113-122. doi: 10.1080/00220485.2017.1285739.

APPENDIX A

We report here below a comparison of the estimated level of skill for selected sub-groups defined by the migrants' origin. A Wald test for the equality of them is also reported.

[TABLE A HERE]

Variable	Metric	Group of respondents	Notes
Level of education	Graded	All	1=lower secondary; 2=upper secondary; 3=tertiary
Highest education achievement of parents	Graded	All	1=lower secondary; 2=upper secondary; 3=tertiary
Language skill in host country	Graded	All	1=basic; 2=intermediate; 3=advanced; 4=mother tongue
Obtained a degree in technical or scientific fields in the last 20 years	Binary	All	0=no; 1=yes
Gender	Binary	All	0=male; 1=female (and vice-versa)
Currently in training or education	Binary	All	0=no; 1=yes
Over-qualified for current job	Binary	Currently working	0=no; 1=yes
Current job is not on fix term base (i.e. it is open ended)	Binary	Currently working	0=no; 1=yes
Current job entails supervision responsibilities	Binary	Currently working	0=no; 1=yes
Best occupation ever reached	Graded	Worked at least once in lifetime	1=elementary occupations; 2=plant and machine operators, and assemblers; 3=craft and related trades workers; 4=skilled agricultural, forestry and fishery workers; 5=service and sales workers; 6=clerical support workers; 7=technicians and associate professionals; 8=professionals; 9=managers

Table 1 Qualitative variables related to the migrants' skill

Variable	Discrimination parameter	Coef.	Std. Err. (p-value)	Difficulty parameter(s)	Coef.	Std. Err. (p-value)
Level of education	α_1	1.896	0.042 (0.000)	$\kappa_{1,1}$	-0.460	0.013 (0.000)
				$\kappa_{1,2}$	0.742	0.015 (0.000)
Highest education achievement of parents	α_2	1.817	0.042 (0.000)	$\kappa_{2,1}$	0.307	0.013 (0.000)
				$\kappa_{2,2}$	1.098	0.019 (0.000)
Language skill in host country	α_3	0.664	0.019 (0.000)	$\kappa_{3,1}$	-2.874	0.078 (0.000)
				$\kappa_{3,2}$	-0.560	0.028 (0.000)
				$\kappa_{3,3}$	1.961	0.055 (0.000)
Obtained a degree in technical or scientific fields in the last 20 years	α_4	1.131	0.048 (0.000)	$\kappa_{4,1}$	3.140	0.101 (0.000)
Currently in training or education	α_5	0.650	0.027 (0.000)	$\kappa_{5,1}$	2.470	0.092 (0.000)
Current job entails supervision responsibilities	α_6	1.427	0.051 (0.000)	$\kappa_{6,1}$	1.617	0.039 (0.000)
Best occupation ever reached	α_2	1.794	0.039 (0.000)	$\kappa_{7,1}$	-0.858	0.018 (0.000)
				$\kappa_{7,2}$	-0.593	0.016 (0.000)
				$\kappa_{7,3}$	-0.110	0.014 (0.000)
				$\kappa_{7,4}$	-0.062	0.014 (0.000)
				$\kappa_{7,5}$	0.756	0.017 (0.000)
				$\kappa_{7,6}$	0.979	0.018 (0.000)
				$\kappa_{7,7}$	1.366	0.022 (0.000)
				$\kappa_{7,8}$	2.468	0.039 (0.000)

Table 2 GRM items parameters estimates.

Country	N	Mean	Std. Err.	Test of equality against													
				CH	CZ	UK	NO	SE	HU	AT	FI	BE	PT	FR	ES	GR	IT
CH	2,681	0.412	0.022		-	**	***	**	-	***	***	***	***	***	***	***	***
CZ	380	0.365	0.042	-		-	-	-	-	***	***	***	***	***	***	***	***
UK	5,027	0.345	0.013	**	-		-	-	-	***	***	***	***	***	***	***	***
NO	1,081	0.315	0.024	***	-	-		-	-	***	***	***	***	***	***	***	***
SE	684	0.309	0.039	**	-	-	-		-	***	***	***	***	***	***	***	***
HU	164	0.303	0.067	-	-	-	-	-		**	***	***	***	***	***	***	***
AT	2,788	0.142	0.018	***	***	***	***	***	**		-	***	***	***	***	***	***
FI	929	0.088	0.034	***	***	***	***	***	***	-		**	***	**	***	***	***
BE	1,899	0.003	0.021	***	***	***	***	***	***	***	**		-	***	***	***	***
PT	684	-0.048	0.030	***	***	***	***	***	***	***	***	-		-	***	***	***
FR	1,227	-0.082	0.031	***	***	***	***	***	***	***	***	**	-		**	***	***
ES	3,577	-0.171	0.016	***	***	***	***	***	***	***	***	***	***	**		***	***
GR	2,412	-0.367	0.014	***	***	***	***	***	***	***	***	***	***	***	***		-
IT	8,590	-0.371	0.009	***	***	***	***	***	***	***	***	***	***	***	***	-	

Table 3 Average skill level by host country and pairwise adjusted Wald tests among means (**p<0.01; **p<0.05; *p<0.1; - not significant).

Country	N	Mean	Std. Err.	Test of equality against												
				UK	CZ	NO	SE	CH	FI	PT	FR	AT	ES	BE	GR	IT
UK	2,216	0.346	0.019	-	***	***	***	***	***	***	***	***	***	***	***	***
CZ	193	0.262	0.057	-	-	-	***	***	***	***	***	***	***	***	***	***
NO	367	0.181	0.044	***	-	-	**	***	***	***	***	***	***	***	***	***
SE	460	0.174	0.048	***	-	-	**	***	***	***	***	***	***	***	***	***
CH	855	0.054	0.035	***	***	**	**	***	***	***	***	***	***	***	***	***
FI	765	0.011	0.036	***	***	***	***	***	**	**	***	***	***	***	***	***
PT	472	-0.100	0.033	***	***	***	***	***	**	-	*	***	***	***	***	***
FR	809	-0.120	0.038	***	***	***	***	***	**	-	-	***	***	***	***	***
AT	1,453	-0.171	0.024	***	***	***	***	***	***	*	-	***	**	***	***	***
ES	2,351	-0.268	0.019	***	***	***	***	***	***	***	***	***	-	***	***	***
BE	715	-0.271	0.031	***	***	***	***	***	***	***	***	**	-	***	***	***
GR	1,889	-0.426	0.015	***	***	***	***	***	***	***	***	***	***	***	-	***
IT	5,984	-0.448	0.010	***	***	***	***	***	***	***	***	***	***	***	-	***

Table 4 Average skill level of non-western extra-EU migrants by host country and pairwise adjusted Wald tests among means (***p<0.01; **p<0.05; *p<0.1; - not significant).

Country	N	Mean	Std. Err.	Test of equality against								
				CH	NO	FR	SE	AT	UK	BE	IT	FI
CH	104	1.235	0.074	***	***	***	***	***	***	***	***	***
NO	59	-0.188	0.095	***	-	-	-	*	**	*	**	**
FR	51	-0.234	0.089	***	-	-	-	-	-	-	-	-
SE	68	-0.290	0.100	***	-	-	-	-	-	-	-	-
AT	222	-0.326	0.053	***	-	-	-	-	*	-	-	-
UK	174	-0.385	0.056	***	*	-	-	-	-	-	-	-
BE	137	-0.419	0.058	***	**	*	-	*	-	-	-	-
IT	40	-0.447	0.098	***	*	-	-	-	-	-	-	-
FI	73	-0.499	0.099	***	**	**	-	-	-	-	-	-

Table 5 Average skill level of refugee migrants by host country and pairwise adjusted Wald tests among means (***p<0.01; **p<0.05; *p<0.1; - not significant).

Average migrants' skill level by host country

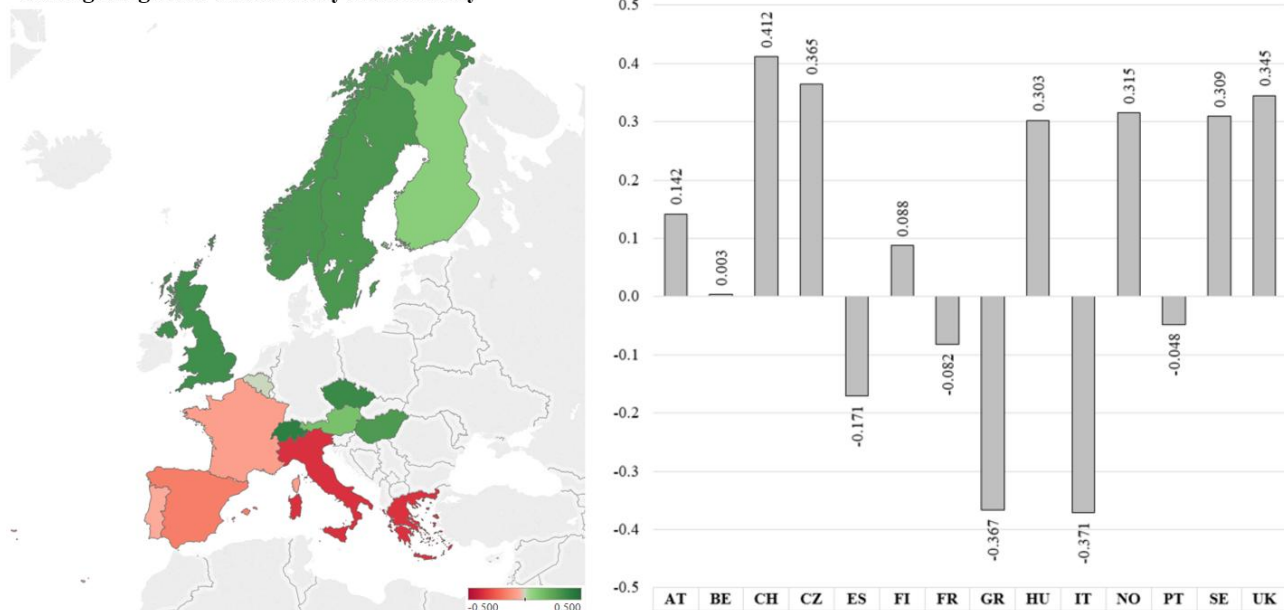


Figure 1 Average migrants' skill level by host country (map and bar graph).

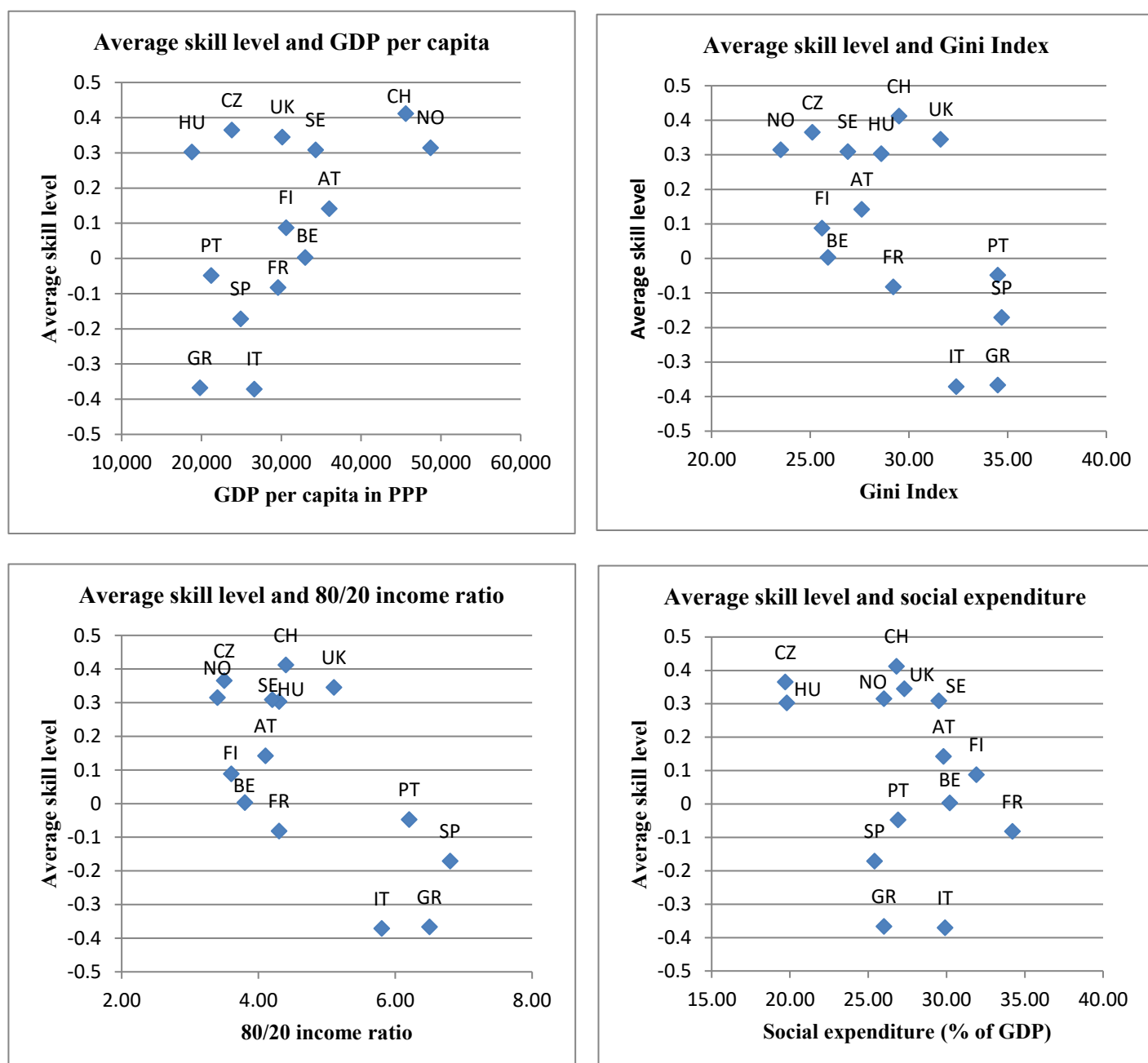


Figure 2 Average skill vs. selected macroeconomic variables in 2014 (Source: our estimations and Eurostat)

Average skill level of non-western extra-EU migrants by host country

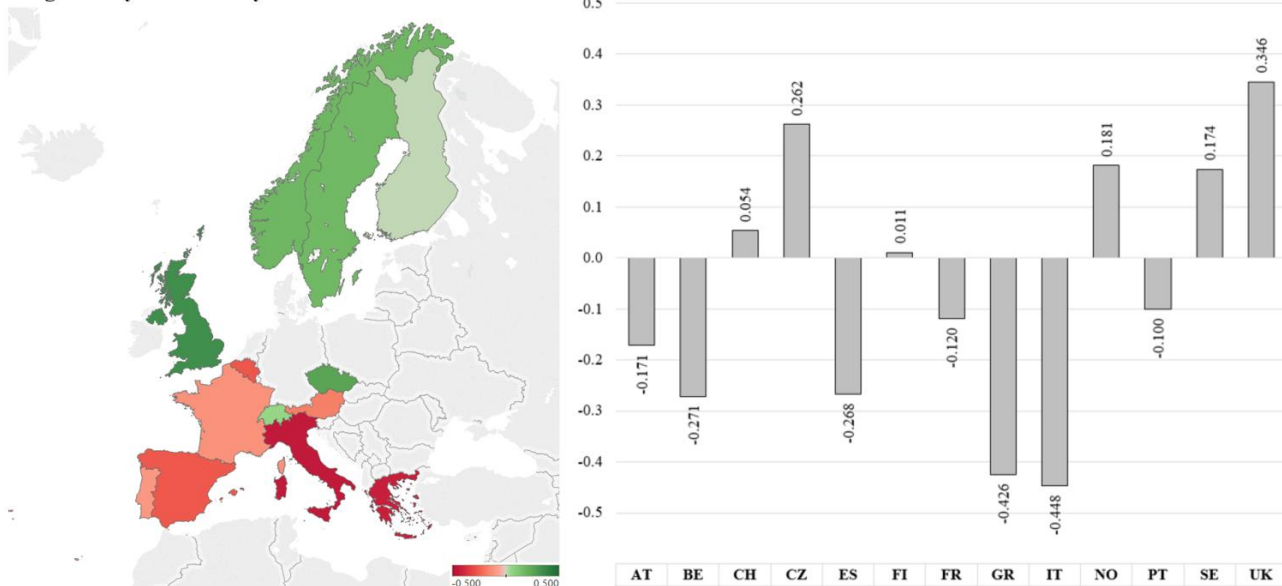


Figure 3 Average skill level of non-western extra-EU migrants by host country (map and bar graph).

Average skill level of refugee migrants by host country

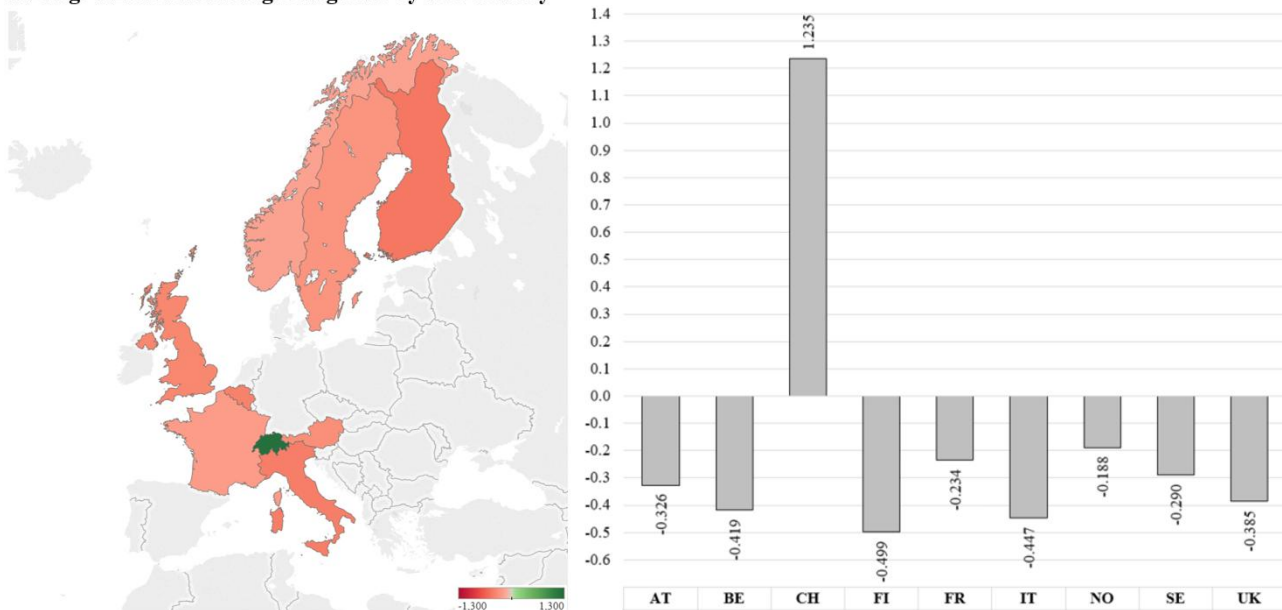


Figure 4 Average skill level of refugee migrants by host country (map and bar graph).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Appendix.docx](#)