Intergenerational mobility and self-selection on unobserved skills: New evidence

Michael Hebsaker, Guido Neidhöfer* and Friedhelm Pfeiffer

Abstract
This study proposes a novel way to examine self-selection on unobserved skills and applies it to a sample of young males seeking asylum in 2015/2016 in Germany. First, the degree of intergenerational mobility of these refugees is assessed, specifically their educational improvement in comparison to their parents’ level of education. Next, the estimates are compared with the level of educational mobility of similar-aged males in the refugees’ regions of origin. The idea is that this difference indicates the pattern of self-selection on unobserved skills such as grit and motivation. Our findings hint at positive selection on such unobserved skills among these young male refugees.

Keywords: Immigrant selection, Asylum seekers, Human capital, Family background

1 Introduction
Between 2014 and 2016 Europe, and in particular Germany, saw a rapid increase in asylum applications. According to Eurostat, in 2016 alone around 750,000 refugees relocated to Germany. Such a significant influx of people has renewed public concerns about the integration of immigrants into society, even influencing voting outcomes in favour of anti-immigration parties (e.g. Dustman et al. 2017; Bratti et al. 2020). Apprehensions regarding refugees and migrants partly stem from cultural differences between migrants and the society of the host country, and partly from the economic burden they may impose on the host country population (Dustmann and Frattini 2014; Fuest 2016, among others). This latter issue is particularly relevant to refugees, since they may begin contributing only in the medium or long term after their assimilation into the labour market is complete; and are often net beneficiaries for a period following their arrival. In addition, the inflow of migrants may adversely affect the labour market opportunities of directly competing residents in the receiving countries (e.g. Borjas and Monras 2017).

Reasoning in the economics of migration literature suggests that migrants are a group of self-selected individuals. Self-selection influences the causes and economic consequences of migration, in particular the labour market achievements of migrants (Borjas 1985; Chiswick 1978). All else equal, higher skilled individuals may assimilate faster into the host country’s society compared to lower skilled individuals. Hereby, the definition of skills is comprised of observable characteristics, such as education, as well as unobserved traits, such as grit, motivation, perseverance or risk preferences.

The Roy-Model of income maximisation predicts that countries with more generous social insurance and benefit systems might attract negatively selected migrants who need more time to assimilate into the host country’s labour market (Borjas 1987, 1999). However, since refugees may differ from economic migrants in several ways, 1Henceforth, we will use the terms asylum seekers and refugees as synonyms.
2 In Germany, for instance, asylum seekers are initially allocated to reception centres where their basic needs are met and are later entitled to receive financial support up to the amount warranted to citizens by the social security system, once they move to private accommodations.
3 The longer-term assimilation of immigrants in the host country can be understood as an intergenerational process where observed as well as unobserved skills of the first generation are transmitted to the second generation of immigrants (Bönke and Neidhöfer 2018; van de Werfhorst and Heath 2019; Zuccotti et al. 2017).
which in turn may affect their economic assimilation, it remains an open question whether and how these considerations apply to refugees. The scant existing international evidence shows contrasting results. Refugees in Sweden and the US have been shown to assimilate faster than other migrants (Luik et al. 2018; Cortes 2004), while the evidence in Norway points to a slower rate of assimilation for refugees (Bratsberg et al. 2014). As a consequence, a lack of information about the relative skill level of current refugees makes it difficult to forecast their economic assimilation prospects within the destination countries (Dustmann et al. 2017). Due to the nature of their displacement—reasons for its occurrence are mostly war, human rights violations or other fatal, unforeseen events—the migration decisions of refugees are presumably less based on economic considerations such as income maximisation. Hence, the selection pattern of refugees may work in two directions; refugees might be more positively or more negatively selected than other migrants from comparable countries of origin.

This study contributes to broaden our knowledge on the skill selection of recently arrived refugees, taking advantage of novel survey data for asylum seekers living near the city of Heidelberg (see Lange and Pfeiffer 2019). The survey includes information on the refugees’ own education background and retrospective questions about their parents’ education. Lange and Pfeiffer (2019) showed that, on average, the young male asylum seekers in this sample and their parents received more years of schooling compared to same-aged males from their country of origin. They thus seem to be a positively selected group (relative to their country of origin) with respect to their amount of time in education. Guichard (2020) confirms these findings for asylum seekers in Germany originating from Iraq and Syria based on a representative survey, and finds a neutral pattern of selection for those that fled from Afghanistan.

In this analysis our aim is to focus on the unobserved skills of the sample of young male asylum seekers from Lange and Pfeiffer (2019) and propose a novel way to measure them. First, we estimate the degree of intergenerational mobility of the refugees in our sample, specifically their educational improvement in comparison to their parents’ level of education. Then, we compare our estimates with cross-country estimates on the level of educational mobility of similar-aged males in the refugees’ regions of origin. In doing so we are able to disentangle the structural component of educational mobility caused by country level characteristics, such as the expansion of schooling or cultural factors. At the same time we are able to keep the largely unobserved component influencing the relative improvement of individuals with respect to their parents. This unobserved component depends mainly on the transmission of personality traits such as motivation, grit, perseverance or the willingness to take risk in the family (see Dohmen et al. 2012; Kosse and Pfeiffer 2012, among others). Hence, the novel idea is that the difference in the degree of intergenerational mobility should indicate the pattern of self-selection on unobserved skills.

Our results show that the refugees in our sample display higher than average relative rates of intergenerational mobility, measured by the association between their own years of schooling and the years of schooling achieved by their parents. Additionally, we estimate the average degree of educational upward mobility, assessed by the probability of refugees to achieve more years of schooling than their parents. We find that refugees from Afghanistan, Iraq, and Sub-Saharan Africa also display higher rates of absolute upward mobility when compared to the same-aged male population in their respective home region. Given our interpretation this finding should be a sign of positive selection on unobservable skills for our sample of refugees.

The remainder of the paper is structured as follows: Sect. 2 briefly summarises the literature on the intergenerational persistence of education. Section 3 provides a detailed description of the data. Section 4 presents and discusses our results. Section 5 concludes.

2 Relative and absolute intergenerational mobility in education

Education, as a proxy for human capital, is strongly determined by family background since education decisions are shaped by parental preferences, the availability of economic resources and credit constraints (Becker and Tomes 1979; Checchi et al. 2013). Hence, parent–child-schooling correlations are strongly related to other measures of social intergenerational mobility such as those based on income or occupation (Blanden 2013; Black and Devereux 2011). An established way to measure the intergenerational mobility of education is to estimate the following linear regression model:

$$S_{i}^{O} = \alpha + \beta S_{i}^{P} + \varepsilon_{i}$$  \hspace{1cm} (1)

where $S_{i}^{O}$ represent the offspring's and $S_{i}^{P}$ the parents’ education in family $i$, measured in years of schooling. The slope coefficient $\beta$ measures the degree of intergenerational persistence (see e.g. Black and Devereux 2011). The higher $\beta$, the stronger the association between parents’ and children's education within the analysed sample. $\alpha$ is a constant, and $\varepsilon_{i}$ an error term.

The cross-country study by Hertz et al. (2007) applies this framework and shows that education is more generationally-persistent in developing countries than in
OECD countries. Furthermore, educational mobility is low in South America and Southeast Asia and rather high in Scandinavian countries. Table S1 in Additional file 1 reports some published estimates of intergenerational mobility for developed countries and for countries in the geographic regions from which the individuals in our sample of refugees originate. These estimates indicate that intergenerational mobility is higher in Germany than in the regions of origin for the refugees in our sample.

The degree of educational mobility within a population is considered to be a summary measure, indicating the persistence of human capital within families over time. As such, it comprises different channels of the human capital production function and intergenerational transmission into one single measure at the cost of losing information about the strength of each component of the relationship. In our analysis, we use the information contained in the mobility estimates to examine the self-selection of refugees on unobserved skills. Hence, we are interested in disentangling the structural component of educational mobility, caused by country level characteristics such as educational expansions or cultural factors, while keeping the unobserved component. According to our argumentation, this unobserved component is the one influencing the relative improvement of individuals with respect to their parent’s education. It depends on personality traits and skills such as motivation, grit, perseverance, or the willingness to take risk. Hence, it should be a suitable indicator of self-selection on unobserved skills.5

We do so by comparing the intergenerational mobility of refugees in our sample with the average degree of intergenerational mobility of same-aged non-migrants in their regions of origin. We retrieve the mobility estimates for the latter from the World Bank’s Global Database on Intergenerational Mobility (GDIM; see Narayan et al. 2018). A higher rate of intergenerational mobility among the subgroup of refugees with respect to their region of origin hints at positive skill selection for this group, while a lower mobility rate indicates that refugees in this group are negatively self-selected.

In this application, we may face a limitation when assessing intergenerational mobility using the estimated slope coefficient in Eq. (1): the slope coefficient \( \beta \) measures the partial correlation between parents’ and children’s years of schooling and, hence, is sensitive to each form of variation within a family from one generation to the next without making a distinction between whether it is an improvement or a deterioration. It is therefore insightful to also estimate an absolute measure of intergenerational upward mobility for our refugee sample, namely, the probability of children having a higher level of education than their parents, given that parents are not in the highest educational category \( m \) (tertiary education6):

\[
Pr(c > p) = Pr(S^O > S^P | \text{age} < m).
\] (2)

The higher this probability, the higher is the average educational upward mobility within the sample. This second indicator of intergenerational mobility, the probability of upward mobility, should be even more insightful as a measure of self-selection with respect to the population of origin, since the first indicator, namely the slope coefficient, captures the degree of regression to the population mean among our sample. To provide a comprehensive description of intergenerational mobility among young male asylum seekers from different countries of origin, both will be reported.

3 Data

The sample of young male asylum seekers that we use stems from the “Real-world Laboratory Survey among Asylum Seekers”, a novel survey data set of asylum seekers who were part of the large influx to Germany in recent years (see Lange and Pfeiffer 2019). The survey contains information on asylum seekers living in two group accommodations close to the city of Heidelberg, in southern Germany. In cooperation with the administration of these group accommodations and the local foreigner’s administration offices, a scientific survey among the asylum seekers was conducted in August/September 2016. Participation in the survey was on a voluntary basis and was open to all individuals over 18 living in the accommodations. The computer-assisted interviews were undertaken by professional and native speaking

---

4 Neidhöfer et al. (2018) show that in most Latin American countries educational mobility has risen substantially for people born in the eighties in comparison to their parents and grandparents.

5 Borjas et al. (2019) use the residuals from earnings regressions to measure the self-selection on unobservable characteristics of Danish emigrants. Corneo and Neidhöfer (2019) measure the unobserved skill level of Italian migrants worldwide, via the probability an individual is either unemployed, or has a high occupational position in the destination country, given their level of education. These two methods for measuring unobserved skill selection are suitable for long-term migrants, but not applicable to this study, since the refugees in our sample were interviewed shortly after arrival, and thus most do not present any measurable labour market outcomes.

6 The necessary years of schooling to be eligible for attending tertiary education tracks vary between most sending countries (usually 12 or 13 regular years of schooling to finish secondary education). We follow the official and country specific UNESCO ISCED Educational Mappings from 2011. See http://uis.unesco.org/en/isced-mappings [latest view 03.06.2019].
interviewers. The design of the survey aimed to cover the main languages spoken by the respondents (Arabic, Dari/Farsi, Tigrinya, Pashtu, English and German).

The data set contains items related to the socio-economic status of the respondents before and after they left their home country. We rely on self-assessment by the respondents in regard to their own and their parents’ educational attainment and follow the literature in employing years of education as a proxy for human capital (e.g. Hertz et al. 2007). Years of schooling is retrieved from the answer to the following question: “How many years did you go to school? (If applicable, including university)” for the asylum seeker, as well as his or her father and mother. It is possible that, particularly among the young refugees in our sample, some individuals might not yet have completed their educational career. Brücker et al. (2016) report that 26% of the refugees dropped out of school or interrupted their education due to consequences of war and flight. Thus, education might be considered as a truncated variable and our intergenerational mobility estimates as a lower bound if individuals were to continue their education in Germany.\footnote{Robustness checks performed on the age range 22–34 are included in the Online Appendix and confirm our main results.}

The legally required minimum age to participate in the survey was 18. To preserve homogeneity within the sample we set the maximum age to 34, excluding the few female respondents (seven observations, less than 2% of the sample), as well as asylum seekers from European countries (eight observations). Hence, from an initial survey sample comprising 370 respondents, we end up with a sample of 206 non-European, male asylum seekers within the age interval of 18–34, with all necessary information available to measure the degree of intergenerational mobility.\footnote{While almost all survey respondents answered the question about their own education (92%), the response rates on fathers’ years of schooling (69%) and mothers’ (72%) are lower. The presence of selective non-response on parental education could potentially bias the intergenerational persistence estimates. However, we find no indication for selective non-response in our sample; response behaviour on parental education is not associated with the average levels of education of individuals. The difference between the average years of schooling of survey respondents that disclosed information about the education of their parents and those who did not is negligible (9.0 vs. 8.98 years of schooling) and statistically undistinguishable from zero. Hence, if non-response on parental education is more likely when the parents have low levels of education, our estimates of intergenerational mobility should not be upwardly biased. Additional file 1: Figure S1 in the Online Appendix shows the distribution of years of schooling of the refugees in our sample and their parents.}

According to the Federal Office for Migration and Refugees (BAMF 2017), 47 percent of the asylum seekers who applied for asylum in 2016 were in the age group 18–34 and more than 70 percent of them were male. Hence, although our sample is not representative of the entire population of newly-arrived asylum seekers in Germany, it focuses on a crucial age interval for young male asylum seekers.\footnote{Brücker et al. (2016) report an average age of 31.2 years within their representative survey among asylum seekers in Germany: A comparison of our data base to the IAB-BAMF-SOEP survey of refugees shows that the asylum seekers in our sample are on average not only younger, but spent also a shorter amount of time in Germany (see Lange and Pfeiffer 2019).}

In the sample asylum seekers from Central Asia account for a total share of 43.5 percent, while 36.5 and 20 percent stem from the Middle East and Africa, respectively.\footnote{Table S2 in the Online Appendix reports the complete list of countries of origin, as well as their respective absolute and relative frequencies. Our sample is not representative of the recent influx of asylum seekers to Germany with respect to the country of origin. According to BAMF (2017), refugees applying for asylum in Germany predominantly stem from Syria (37%), Afghanistan (18%) and Iraq (13%). Asylum seekers from Afghanistan and Gambia are relatively overrepresented in our sample, while Syrian asylum seekers seem to be relatively underrepresented.} Almost 41 per cent of the asylum seekers stem from Afghanistan, 17 percent from Syria and 15.5 percent from Iraq. Gambians constitute the largest group of African asylum seekers in our sample (10.2 percent). The remaining 16.8 percent stem from other Asian and African countries.\footnote{Additional file 1: Table S2 in the Online Appendix reports the complete list of countries of origin, as well as their respective absolute and relative frequencies. Our sample is not representative of the recent influx of asylum seekers to Germany with respect to the country of origin. According to BAMF (2017), refugees applying for asylum in Germany predominantly stem from Syria (37%), Afghanistan (18%) and Iraq (13%). Asylum seekers from Afghanistan and Gambia are relatively overrepresented in our sample, while Syrian asylum seekers seem to be relatively underrepresented.} The figure is in line with results of other studies investigating educational patterns of newly-arrived asylum seekers. Buber-Ennser et al. (2016) for Austria, as well as Brücker et al. (2016) for Germany, find comparable average years of schooling based on survey data. None of these studies report the educational attainment of parents.

Per the literature, we use the maximum level of education among both parents as a proxy for parental education in order to estimate intergenerational mobility.\footnote{Other studies, such as Hertz et al. (2007), use the average among both parents instead. As shown by Neidhöfer et al. (2018), country rankings usually do not change when applying one or the other variable for sufficiently high and similar levels of spouse correlations among parents across countries. The correlation between father’s and mother’s years of schooling in our sample is 0.65, ranging from 0.48 for refugees from Sub-Saharan Africa to 0.74 for refugees from MENA countries.} In our sample, we observe on average 5.97 years of education for fathers and 3.85 years for mothers. Decile values reveal a clustering pattern; particularly in the case of mothers (57% with zero years of schooling) and fathers (38.5% with zero years of schooling), but less so for the
individual’s level of educational attainment (13% have zero years of schooling).

4 Results
4.1 The intergenerational mobility of male asylum seekers
Table 2 shows the estimates for the $\beta$ coefficient in Eq. (1), further controlling for age and including country of origin fixed effects. The dependent variable is the years of schooling of the respondent. Column (1) and (2) show our main specifications, measuring parental human capital by the maximum level of schooling among both parents. We find a point estimate of 0.36 for the average degree of intergenerational persistence measured by the slope coefficient within our sample of young male asylum seekers. This means that an additional year of parental schooling is associated with an increase of about one third of a year of schooling for the next generation. In the subsequent columns we measure the association with father’s and mother’s years of schooling separately, and then include both in the regression. The estimates do not change the overall pattern substantially and show that the father’s education is a better predictor of children’s education than the education of the mother. As is evident, the inclusion of age does not change the estimates significantly.

Additional file 1: Table S1 in the Online Appendix surveys part of the literature on intergenerational mobility estimates. We include the most comparable ones to our estimates, relying on the same methodology, similar age cohorts and parent-son pairs (see Additional file 1: Table S1 for additional information on samples and

| Table 1 Descriptive Statistics on Individual and Parental Years of Schooling |
|-----------------------------------------------|
| $S_i^O$: Individual years of schooling of asylum seekers; $S_F^O$: Father’s years of schooling; $S_M^O$: Mother’s years of schooling; $S_P^O$: Years of schooling of the parent with the highest years of schooling among the two. Last column displays the average and distribution of the age of respondents. Source: Own estimates. Sample taken from ‘Real-world Laboratory Survey among Asylum Seekers’ |
| Variable          | $S_i^O$ | $S_F^O$ | $S_M^O$ | $S_P^O$ | Age        |
|-------------------|---------|---------|---------|---------|------------|
| Mean              | 9.00    | 5.97    | 3.85    | 6.48    | 23.34      |
| Standard deviation| 4.89    | 5.87    | 5.21    | 5.91    | 3.92       |
| 10th decile       | 0       | 0       | 0       | 0       | 19         |
| 20th decile       | 4       | 0       | 0       | 0       | 20         |
| 30th decile       | 7       | 0       | 0       | 0       | 20         |
| 40th decile       | 9       | 2       | 0       | 3       | 21         |
| 50th decile       | 11      | 5.5     | 0       | 6       | 22         |
| 60th decile       | 12      | 8       | 3       | 9       | 24         |
| 70th decile       | 12      | 10      | 6       | 12      | 26         |
| 80th decile       | 12      | 12      | 9       | 12      | 27         |
| 90th decile       | 14      | 14      | 12      | 14      | 29         |

| Table 2 Slope coefficient of parent–child associations in years of schooling |
|---------------------------------------------------------------|
| (1) (2) (3) (4) (5) (6) |
|-----------------------------------------------|
| Schooling parents ($S_P^O$)       | 0.35*** (0.05) | 0.36*** (0.05) | – | – | – | – |
| Schooling father ($S_F^O$)        | – | – | 0.33*** (0.05) | – | 0.27*** (0.06) | 0.27*** (0.06) |
| Schooling mother ($S_M^O$)        | – | – | – | 0.29*** (0.05) | 0.11 (0.07) | 0.11* (0.07) |
| Age                             | – | 0.21** (0.08) | – | – | – | 0.20** (0.08) |
| Constant                        | 5.71*** (0.63) | 1.03 (1.87) | 5.94*** (0.61) | 6.83*** (0.59) | 5.96*** (0.61) | 1.45 (1.90) |
| Adj. $R^2$                      | 0.204 | 0.241 | 0.207 | 0.156 | 0.211 | 0.232 |

Sample comprises 206 individuals. Schooling Parents refers to the years of schooling of the parent with the highest years of schooling among the two. Schooling Father/Mother refers to the years of schooling. Regressions controlling for country of origin fixed effects. Robust standard errors in parentheses. Statistical significance level of the estimates: * for $p < .10$, ** for $p < .05$, and *** for $p < .01$. Source: Own estimates, sample taken from ‘Real-world Laboratory Survey among Asylum Seekers’.
descriptive statistics). The comparison shows that individuals in our sample of refugees are, on average, more mobile than the population in most transitioning and developing countries, but less mobile than the population in most refugee reception countries, such as Germany.

5 Comparison to region of origin
Lange and Pfeiffer (2019) uncovered a positive pattern of selection on years of schooling for the refugees in our sample. In addition, their results show that the parents of the refugees in our sample have, on average, higher educational achievements than the population of parents in the country of origin. Hence, a priori no conclusions can be drawn from these findings about the degree of intergenerational mobility of the asylum seekers in our sample; it might be higher, lower or the same in comparison to the average intergenerational mobility of their peers in their country of origin.

In this section we compare the intergenerational mobility estimates for our sample of asylum seekers with the overall level of intergenerational mobility in their country or region of origin. For this purpose, we estimate the intergenerational mobility separately for subgroups of refugees, clustered by their country or region of origin. We compare these assessments with estimates retrieved from the World Bank's Global Database on Intergenerational Mobility (GDIM; 2018; see also Narayan et al. 2018). Intergenerational persistence estimates retrieved from the GDIM pertain to even-aged males belonging to the most comparable birth cohort to the refugees in our sample, namely individuals born between 1980 and 1989.13 The refugees in our sample were born between 1982 and 1998, and are slightly younger than the cohort chosen from the GDIM for comparison purposes. The latter, however, is the last available cohort in the database. Further reducing our sample of refugees limits the power of the analysis due to the small sample size. However, our intergenerational persistence estimates obtained with a restricted sample of refugees aged 22–34, i.e. born between 1982 and 1994, are consistent with our main estimates (see Additional file 1: Tables S3 and S4 in the Online Appendix). For this subsample of on average older asylum seekers we observe even higher levels of educational upward mobility.

For both Afghanistan and Iraq, we have enough observations to estimate the intergenerational mobility indices within our sample and make a direct comparison, as the countries are included in the GDIM. For Syria, since no country specific estimates are available in the GDIM, we built a synthetic comparison group based on the same cultural and geographic region and income group the countries are ranked in (middle income countries).14 We then report the unweighted average among the countries belonging to this group.15 To provide further comparison groups with sufficiently high numbers of observations, we form two further subgroups of refugees by their region of origin—MENA and Sub-Saharan Africa—and compare the estimates with the unweighted average of all country estimates contained in the GDIM in these regions (see Table 3 notes).

Table 3 shows our estimates and the World Bank estimates. We report the regression-based index for relative mobility, obtained by estimating Eq. (1) on our own sample and retrieved from the World Bank data for each country or region of origin, as $\beta$. The table also contains the predicted probability that individuals attained a higher level of education than their parents, given that neither of the parents is in the highest education category.

Iraq (32), Algeria (3); Sub-Saharan Africa: Eritrea (14), Gabon (1), Gambia (21), Niger (1), Nigeria (1). Regional composition in the GDIM: MENA: Djibouti, Egypt, Iran, Iraq, Jordan, Lebanon, Morocco, Tunisia, West Bank and Gaza, Yemen; Sub-Saharan Africa: Benin, Burkina Faso, Central African Republic, Comoros, Ethiopia, Guinea, Guinea-Bissau, Liberia, Madagascar, Mali, Mozambique, Malawi, Niger, Rwanda, Senegal, Sierra Leone, South Sudan, Chad, Togo, Tanzania, Uganda. Source: Own estimates, sample taken from `Real-world Laboratory Survey among Asylum Seekers'; and World Bank Global Database on Intergenerational Mobility (GDIM).

The average probability that individuals in the sample improve their level of education with respect to their parents is 0.6, which is consistent with the high degree of relative mobility estimated within this sample. Excluding individuals younger than 22, the probability of upward mobility is even higher and amounts to 0.67. Hence, the

---

13 The comparison data selected from the GDIM builds on large Microdata retrieved from population surveys, which were conducted in the corresponding countries. Selected data is restricted to parent-son pairs given that the child is born in the 1980’s birth cohort. The underlying population surveys were conducted between 2010 and 2016.

14 The synthetic comparison group for Syria drawn from the GDIM includes the following countries: Djibouti, Egypt, Morocco, Tunisia, West Bank and Gaza, and Yemen. Country classifications are drawn from published World Bank country classifications, online: https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups. [10.04.2019].

15 The other alternative would be to build the average by weighting countries using their population size. However, this would lead to a bias towards the intergenerational persistence estimate of more populous countries within the region. Instead, our aim is to evaluate the intergenerational mobility of the asylum seekers in our sample with respect to the level of mobility of residents in their regions of origin as a benchmark. For this purpose, the unweighted average over all countries in the region yields a more informative measure.
Table 3 Educational Mobility of Asylum Seekers in Comparison to their Region of Origin

| Country/Region | N  | Slope coefficient (β) | Absolute upward mobility (Pr(c > p)) | Pr(c > p) GDIM |
|----------------|----|-----------------------|--------------------------------------|----------------|
| Afghanistan    | 84 | 0.42 (0.09) 0.24–0.60  | 0.53 (0.06) 0.41–0.64                | 0.36           |
| Syria          | 35 | 0.23 (0.12) 0.00–0.45  | 0.62 (0.10) 0.45–0.80                | 0.66           |
| Iraq           | 32 | 0.31 (0.11) 0.10–0.52  | 0.54 (0.10) 0.34–0.94                | 0.45           |
| MENA           | 78 | 0.25 (0.08) 0.10–0.40  | 0.59 (0.06) 0.46–0.71                | 0.64           |
| Sub-Saharan    | 38 | 0.45 (0.10) 0.26–0.64  | 0.74 (0.07) 0.61–0.88                | 0.35           |
| Total Sample   | 206| 0.35 (0.05) –        | 0.60 (0.04) –                         | –              |

Estimates pertain to the respective subgroups of young male refugees in our sample by their country or region of origin. β is the slope coefficient retrieved from the estimates of Eq. (1). Pr(c > p) is the probability of children attaining higher education than their parents, given that the parents are not in the highest category (tertiary education). Below our estimates in parentheses are the bootstrapped standard errors and 95% confidence intervals obtained with 1000 replications. GDIM: World Bank estimates retrieved from the GDIM for males of the 1980’s birth cohort. Point estimates shown for single countries, while average values for synthetic control groups include the range [min–max] in brackets. N shows the number of observations within the subgroup. Regional composition in our sample: MENA: State of Palestine (6 observations), Syria (35), Afghanistan (2).

regression-based estimate for the sample of young male refugees is not driven by a downward mobility pattern. Figure 1 visualises the difference between the degree of intergenerational mobility of male asylum seekers measure in our sample and the World Bank estimates for comparable peers in the country or region of origin. Here, for reason of simplification, intergenerational mobility measured by the slope coefficient is displayed as 1 − β.

The estimates show that the refugees in our sample display consistently higher rates of intergenerational mobility than the average for their respective country or region of origin; with the exception of refugees from Syria and the MENA region, whose upward mobility rate is similar to the average rate of comparable individuals in their region of origin. For Syria and the MENA region the set of countries in our sample differs from the ones included in the World Bank data, and hence may lead to imprecise results regarding the comparison of these two groups. Furthermore, because of our small sample size and since the GDIM does not provide standard errors of the point estimates, the validity of statistical tests for differences between these estimates remains limited.

6 Conclusions

In this study we analysed the self-selection of a sample of recently arrived male asylum seekers near the city of Heidelberg based on a novel method. Previously, Lange and Pfeiffer (2019) showed that these young male asylum seekers are positively selected on observed years of schooling compared to their country of origin. To assess the degree of self-selection on unobserved characteristics for these young male refugees, this study measured their degree of intergenerational mobility and compared it to the average for comparable individuals in their region of origin. We found that the refugees in our sample display, on average, a higher degree of intergenerational mobility than a same-aged reference group in their country of origin. In conclusion, the findings indicate that the asylum seekers in our sample presumably are likely not a negatively, but a positively selected group in comparison to the resident population in their regions of origin.

Our sample covers young male refugees living in two group accommodations in a German municipality. This group of young male asylum seekers forms a major component of refugees in Germany, and the findings should therefore be of particular importance. The quota-based and random allocation of refugees among the German Federal States and municipalities may support the external validity of these results, although this needs further
confirmation with larger datasets. Because of the relatively small sample size, the analysis has only limited statistical power. Although our sample and estimates should be comparable with the World Bank estimates, to some degree, uncertainty persists. The measurement of educational achievement in our data is the years of schooling as indicated by the respondents, while the World Bank analysis uses the imputed regular years of schooling, based on the educational degree obtained as indicated by the respondents. Measurement error might challenge the comparison of estimates. Furthermore, it goes beyond the scope of this work to evaluate the relative skill level, and associated integration prospects, of certain refugee groups in comparison to each other.

Our analysis adds new and relevant, although preliminary, insights regarding the economic implications of refugee inflows for recipient countries. As shown by Guichard (2020) the bulk of recently arrived refugees in Germany are positively self-selected, measured by their level of schooling, which is also the case for the individuals in our sample. Brückner et al. (2019, 2020) show that refugees’ German language skills improved after one year, as well as their participation in the labour market and investments in education and training. Our findings suggest that young male asylum-seekers may also be positively self-selected on unobserved skills, such as motivation, grit, perseverance, or willingness to take risk. If the unobserved skill level of refugees is relatively high, as our interpretation of the findings suggests, this may have affected their assimilation prospects positively thus far. Future research could examine whether our findings can be generalized for larger groups of asylum seekers.

Supplementary Information
The online version contains supplementary material available at https://doi.org/10.1186/s12651-021-00291-3.

Acknowledgements
The data utilized in this study has been collected during a research project which is a part of “Real-World Laboratory: Asylum Seekers in the Rhine-Neckar-Region”, supported by the State Ministry of Science, Research and the Arts of Baden-Württemberg and the Humboldt Foundation. The views expressed in this study are those of the authors, and do not necessarily reflect the views of the State Ministry. We gratefully acknowledge insightful comments to an earlier version from Sara Ben Yahmed, Sarah McNamara, Martin Lange, and two anonymous referees, which helped to improve the paper. Remaining errors are solely our own.

Authors’ contributions
The authors contributed equally to the analysis and the writing of the article. All authors read and approved the final manuscript.

Funding
This study was funded by Ministry of Science, Research and the Arts of Baden-Württemberg under grant number 8809–12/206/1.

Availability of data and materials
The data have been collected by the authors during a research project which is a part of “Real-World Laboratory: Asylum Seekers in the Rhine-Neckar-Region” and can be accessed onsite at the ZEW—Leibniz Centre for European Economic Research in Mannheim.

Declarations
Competing interests
The authors declare that they have no competing interests.

Received: 30 September 2020 Accepted: 9 March 2021
Published online: 22 March 2021

References
Assaad, R., Saleh, M.: Does improved local supply of schooling enhance intergenerational mobility in education? evidence from jordan. World Bank Econ. Rev 32(3), 633–655 (2018)
Azam, M., Bhatt, V.: Like father, like son? intergenerational education mobility in India. Demography 52(6), 1929–1959 (2015)
Azomahou, T., Yitbarek, E.A. (2016). Intergenerational education mobility in Africa: has progress been inclusive? Policy Research Working Paper 7843
BAMF: Das Bundesamt in Zahlen 2016 – Modul Asyl. Bundesamt für Migration und Flüchtlinge, Berlin (2017)
Becker, G.S., Tomes, N.: An equilibrium theory of the distribution of income and intergenerational mobility. J. Political Economy 87(6), 1153–1189 (1979)
Binzel, C., Carvalho, J.P.: Education, social mobility and religious movements: the islamic revival in Egypt. Econ. J. 127(607), 2553–2580 (2017)
Black, S., Devereux P.J. (2011). Recent developments in intergenerational mobility, in Orley Ashenfelter and David Card, eds., Handbook of Labor Economics, Volume 4B, 1487–1541.
Blanden, J.: Cross-country rankings in intergenerational mobility: a comparison of approaches from economics and sociology. J. Econ Surveys 27(1), 38–73 (2013)
Bönke, T., Neidhöfer, G.: Parental background matters: Intergenerational mobility and assimilation of italian immigrants in Germany. German Econ. Rev 19(1), 1–31 (2018)
Borjas, G.: Assimilation, changes in cohort quality, and the earnings of immigrants. J. Labor Econ. 3(4), 463–489 (1985)
Borjas, G.: Self-selection and the earnings of immigrants. American Econ. Rev 77(4), 531–553 (1987)
Borjas, G: Immigration and Welfare Magnets. J Labor Econ. 17(4), 607–637 (1999)
Borjas, GJ, Monras, J.: The labour market consequences of refugee supply shocks. Econ. Policy 32(91), 361–413 (2017)
Borjas, GJ, Kauppinnen, I., Pouvtvara, P.: Self-selection of emigrants: theory and evidence on stochastic dominance in observable and unobservable characteristics. Econ. J. 129(617), 143–171 (2019)
Bratsberg, B., Raum, O., Raad, K.: Immigrants, Labour Market Performance and Social Insurance. Econ. J. 124(590), 644–683 (2014)
Bratti, M., Deiana, C., Havari, E., Mazzarella, G., Meroni, E.C.: Geographical proximity to refugee reception centres and voting. J Urban Econ. 120, 103290 (2020)
Brücker, H., Croiser, J., Kosyakova, Y., Kröger, H., Pietrantuono, G., Rother, N., Schupp, J.: Language skills and employment rate of refugees in Germany improving with time. DIW Weekly Report 9(4/6), 49–61 (2019)
Brücker, H., Kosyakova, Y., & Schüßl, E. (2020). Fünf Jahre seit der Fluchtmigration 2015: Integration in Arbeitsmarkt und Bildungssystem macht weitere Fortschritte (No. 4/2020). IAB-Kurzbericht.

Additional file 1: Table S1. Studies on the intergenerational Persistence of Years of Schooling. Table S2. Number of observations by country of origin. Figure S1. Distribution of educational attainment among asylum seekers in our sample. Table S3. Robustness check-Slope coefficient of parent-child associations in years of schooling (age range 22–34). Table S4. Robustness check-Educational Mobility of Asylum Seekers in Comparison to their Region of Origin (age range 22-34).
