Centre of Excellence
for Science and Innovation Studies

CESIS Electronic Working Paper Series

Paper No. 485

Productivity of refugee workers and implications for innovation and growth

Christopher F Baum

Hans Lööf

Andreas Stephan

Klaus F. Zimmermann

March, 2022

The Royal Institute of technology
Centre of Excellence for Science and Innovation Studies (CESIS)
http://www.cesis.se

Productivity of refugee workers and implications for innovation and growth*

Christopher F Baum (Boston College, DIW Berlin and CESIS)

Hans Lööf† (Royal Institute of Technology, Stockholm)

Andreas Stephan (Linnaeus University and DIW Berlin)

Klaus F. Zimmermann (UNU-MERIT, Maastricht University, CEPR and GLO)

March 24, 2022

Abstract

Occupational sorting, classified by the skill-biased technical change theory, explains the largest share of the estimated wage variation of native and refugee immigrant workers. Refugee workers are less likely to be employed in high-paid jobs and more likely to be sorted into low-skilled jobs than comparable native-born workers. Within most occupations, the differences are small or non-existent. In several STEM occupations, commonly regarded as strategic for innovation-driven economies and in which many companies face difficulties in recruiting personnel, the gap is modest or even reversed. Considering wages as a proxy for productivity, this paper using Swedish register data has implications for innovation and growth in many OECD countries characterized by an aging population and shortages of skilled workers.

JEL: C23, F22, J24, J6, O15

Keywords: Blinder–Oaxaca decomposition, employer-employee data, occupational sorting, productivity, refugee immigrants

---

*We thank two anonymous reviewers and participants at the following seminars and conferences for comments and suggestions on earlier versions of the paper: Portuguese Stata Conference 2020, Porto; University of Minho, Braga, Swedish Ministry of Employment 2020, Stockholm; Institute for Evaluation of Labor Market and Educational Policy 2020, Uppsala, United Nations University, Maastricht; Economic and Social Research Institute on Innovation and Technology (MERIT) 2020, Maastricht; the International Economic Association (IEA) World Congress 2021, and University of Birmingham 2022

†Corresponding author: hans.loof@indek.kth.se Author contact emails: baum@bc.edu, Andreas.Stephan@lnu.se, klaus.f.zimmermann@gmail.com

1 Introduction

Many OECD economies are simultaneously experiencing an increasing share of old age in the population and a shortage of workers in the labor market. As evolution of the demographic structure is a low-frequency phenomenon, immigration may be a strategic factor for these economies. Currently, refugees are a main source of immigration in many industrialized countries. The number of refugee immigrants is expected to remain high over the coming decades due to conflicts and environmental disasters.[1]

This paper examines whether refugee immigrants can alleviate the negative impacts of skilled labor shortages on innovation, productivity, and growth by accumulating necessary skills for both manual and cognitive tasks in a knowedge-based economy.[2]

The empirical analysis is conducted on the Swedish labor market for several reasons. Sweden is one of the Western world's largest refugee recipients in both relative and absolute terms.[3] It has administrative register data that covers the entire population of individuals and firms. The linked employer-employee data allows us to observe unique firms as well as unique individuals over time whether employed or unemployed. Sweden is also ranked as a top nation in

[1] More than 80 million people around the world have been forced to flee their homes. Nearly one third of them are classified as refugees. This number is the highest ever seen. Around half of the forced migrants are under the age of 18. About 7 million refugees have sought protection in OECD countries.

[2] According to Kuznets (1960), changes in age structure can affect the medium- and long-term macroeconomic prospects (Kuznets cycles) since age groups differ in their (i) savings behavior, according to the life-cycle hypothesis; (ii) productivity levels, according to the age profile of wages; (iii) labor input, as the young and old tend not to work; (iv) contribution to innovation, with young and middle-age workers contributing the most; and (v) investment opportunities, as firms target the different needs and increasing share of old age in the population.

[3] https://www.unhcr.org/5d08d7ee7.pdf

innovation performance.[4]

To study the labor market performance of refugee workers, we focus on wage earnings as a proxy for labor productivity, assuming that the price of labor is determined by marginal productivity in accordance with the marginal productivity theory of wages.

Although previous studies have investigated wage differences between native workers and refugee-immigrant workers and report substantial disparities, there is little systematic evidence whether this is a "Mandelbrot's fractal phenomenon" (Griliches and Mairesse, 1995) that remains when comparing aggregates, such as the entire labor market or total manufacturing, to data classified in disaggregate form. We fill this gap by examining wage performance within occupations and at the work-task level.

The dramatic increase in people seeking protection in Western countries has provoked lively debates in the economics literature about refugees' impact on the labor markets (Balkan and Tumen, 2016; Borjas and Monras, 2017; Card, 1990; Peri and Yasenov, 2019; Clemens and Hunt, 2019; Foged and Peri, 2016; Tumen, 2016) as refugee immigration has been a major policy issue in almost all OECD countries. There is a widespread perception that refugee flows, usually from low-income countries, consist of large masses of unskilled laborers which put a strain on national economies. This statement is contradicted by recent research arguing that well-educated and highly skilled people are more likely to be immigrants than people with less education and skills: see, for instance, Grogger and Hanson (2011) and Peri (2016).

The largest body of research on labor market performance of refugee im-

[4] https://ec.europa.eu/docsroom/documents/46013

migrants concerns employment and unemployment. The relative wage level, as an indicator of productivity, is still an understudied area (Fasani, Frattini and Minale, 2018). A primary problem concerns the availability of data, as discussed by Cortes (2004), Chin and Cortes (2015), Evans and Fitzgerald (2017), and Dustmann and Görlach (2016). Prevailing wage studies are generally based on survey data limited to a single cohort or a few cohorts of migrants, and they are often conducted without a distinction between voluntary and forced migrants.

The existing studies on refugee integration following the wage performance of unique individuals in a longitudinal dimension is limited to only a few analyses using data from North America or northern Europe. Reviewing the scarce literature exploiting administrative register data or repeated surveys on unique individuals, Brell, Dustmann and Preston (2020) report that refugees' wage earnings ten years after migration as a fraction of the mean wages of natives, conditional on being in employment, are about 60% in Canada, United States and Finland, and about 75% in Norway and Sweden.

To analyze these differentials within components of the labor force, we adopt the occupational classification scheme of the skill biased technical change (SBTC) literature based on Autor, Levy and Murnane (2003), Acemoglu and Autor (2011), and Acemoglu and Restrepo (2018). This literature highlights the increasing wage gap between non-routine and routine tasks and, in particular, between cognitive and manual work tasks as a consequence of technical change and increased skill intensity.

While technical change traditionally has been viewed as factor-neutral, the SBTC approach builds on the idea that new technologies, changes in production processes, and changes in the organization of work are more complementary to skilled workers. As this technological shift increases the relative productivity of skilled workers, it tends to increase demand for skilled workers and decrease demand for low and unskilled workers, altering relative wages (Murnane, Willett and Levy, 1995). If refugee workers have a larger likelihood to be sorted into low-skilled jobs than native-born workers with a similar skill background, this may contribute to the wage gap despite controlling for individual characteristics.

The data are provided by Statistics Sweden and contain extensive information on all individuals living in the country born between 1954–1980 as well as employee data linked to employer data. We consider the labor market performance over the period 2003–2013 for native-born workers and for refugee immigrants, who arrived before 1997. A refugee is defined as an asylum seeker whose request for refugee status has been approved and therefore has full access to the labor market.[5]

Three groups of immigrants are analyzed: European refugees arriving in the 1990s, non-European refugees entering Sweden in the 1990s, and all pre-1990 refugee immigrants. We make this distinction in order to see whether cultural distance matters for the labor market integration of refugees, and also to include a refugee group that arrived in Sweden before the early 1990s, as the more recent group is mainly from the former Yugoslavia.

Extensive research from different disciplines suggests that refugees on arrival are disadvantaged in social and economic terms relative to the native

[5]In accordance with the framework for the international regime of refugee protection. See https://www.unhcr.org/3d4aba564.pdf

population, and that several problems tend to persist. The literature lists several factors which may contribute to the disadvantage vis-à-vis native citizens, e.g., host-country's applicable human capital, including language and job skills (De Vroome and Van Tubergen, 2010), recognition of credentials for qualifications and previous work experience (Ager and Strang, 2008), initial employment bans for asylum seekers (Marbach, Hainmueller and Hangartner, 2018), levels of schooling (Chin and Cortes, 2015), time in the country and experience (Bevelander, 2020), residential area (Connor, 2010), social networks (Auer, 2018), uncertainties about duration of stay (Schock, Böttche, Rosner, Wenk-Ansohn and Knaevelsrud, 2016), and physical and mental health conditions related to incidents before arrival to the host country and discrimination (Ruiz and Vargas-Silva, 2018). Drawing upon this literature, we include indicators and data on host-country applicable human capital, education level, time in the country, professional experience, and area of residence in the analysis.

Our data have several restrictions. First, we exclude self-employed workers assuming that they are not obviously comparable with employed workers. Second, we focus on individuals born between 1954 and 1980. Thus, we compare wage levels for workers aged from 33 to 59 years. Third, we only study refugee immigrants arriving before 1997. Forth, the empirical analysis is conducted for 'established' workers, defined as those earning at least 60% of the median monthly wage.[6]

[6] The latter restriction is in accordance with Statistics Sweden, which separates the labor force into six categories: established position, insecure position, weak position, university studies, other studies, and neither working nor studying. The threshold for being classified into an established position corresponds to about 60% of the median wage in the labor market (61% in the year 2019). Over the period 2003–2013, 84% of matched natives in our study were classified as established in the labor market, compared to 72% of European refugees, 60% of non-European refugees, and 65% of pre-1990 refugees.

The empirical approach consists of the following steps. First, we employ a coarsened exact matching (CEM) approach where a control group of native-born individuals from the full population is chosen for having the same characteristics as the refugee immigrants. Those characteristics include age, gender, marital status, number of children, education, and place of living. We then estimate correlated random effects (CRE) models and apply these results in two decomposition approaches. Second, we estimate a wage equation by using the correlated random effects panel approach (Mundlak, 1978; Wooldridge, 2010). This approach allows us to control for unobserved heterogeneity at the individual level while including the effects of time-invariant regressors such as group membership.[7] In a third step, we apply the Owen–Shapely value decomposition of explanatory factors to the CRE estimate to explain wage variation in the empirical model. Finally, based on the wage-earnings equations, we apply the Blinder–Oaxaca technique (Blinder, 1973; Oaxaca, 1973) to decompose observed differences in wage earnings into explained and unexplained components.

Our estimates confirm previous studies showing a large overall wage divergence (22.6%) between established native and refugee immigrant workers. Applying a decomposition approach, we are able to explain 20.5% of that difference: almost the entire gap. Occupational sorting into work tasks, as classified by the SBTC theory, accounts for the largest share of the observed wage variation. In occupations with routine and manual work tasks, which account for more than half of the Swedish labor market, the wages between refugees and comparable natives tend to converge over time. However, in more skill-intensive occupations, there is an average unexplained difference for refugee

[7] In a robustness test, we apply several IV approaches and account for selectivity bias.

workers, whose wages are more than 10% lower. We also show that the results vary within these occupations. In several STEM (science, technology, engineering and mathematics) occupations, commonly regarded as strategic in innovation-driven economies and subject to skilled labor shortages, we find only marginal wage differences, or even higher wages for refugee workers.

Using wages as a proxy for labor productivity, our study makes an important contribution by suggesting that refugee immigrants may play an important role in remedying labor shortages in all parts of the labor market. The results also indicate that this potential has so far been utilized primarily in low-wage occupations. A policy conclusion is that the recruitment of skilled refugee workers into more knowledge-intensive, high-wage occupations should be encouraged to enhance productivity and economic growth. Providing insights into labor market integration from previous waves of refugees, we think that our findings are relevant also for current and future refugee crises.[8]

The rest of the paper is structured as follows. Section 2 provides details of the data and their descriptive statistics. Section 3 describes the empirical strategy. Empirical results are provided in Section 4. Section 5 discusses robustness tests, and Section 6 concludes.[9]

[8] Already during the first week of war in Ukraine, the UNHCR estimates up to 5 million fleeing people if Russia would occupy the country. By comparison, 2.3 million people fled their homes between 1989 and 1992 as a result of the collapse of the six republics of Yugoslavia.

[9] Our online appendix provides information on data and presents equations, estimates, and figures not reported in the paper.

2 Data and Descriptive Statistics

2.1 Data

We use employer-employee register data provided by Statistics Sweden. Non-refugee immigrants are not included in the analysis. The self-employed are also excluded as they exhibit quite different behavior than the wage earners who are the focus of our analysis.

Our sample contains extensive information on all individuals in Sweden born between 1954 and 1980 as well as variables related to all firms in Sweden, accessed through the remote MONA (microdata online access) delivery system. The variables constructed from these sources include population groups (natives, various refugee groups), demographics (gender, age, marital status, preschool children), education, citizenship, work characteristics (occupational tasks, work experience, wage), firm characteristics (industry, firm size) and geography (municipalities, rural areas, regions).

The key variables are defined in Table 1. We use information on the migration background of a person to identify all refugee immigrants who arrived in Sweden before 1997 and were granted asylum. They are separated into three refugee groups: (1) those from European countries arriving during the period 1990–1996, (2) those from non-European countries arriving during the same period, and (3) those arriving in Sweden between 1980–1989 without classifying their country of origin. We split the first two groups because one could assume that European refugee immigrants may be subject to less discrimination in the labor market than non-European refugees Lundborg (2013). This intuition is also supported by the descriptive statistics reported in Table 2 showing that European refugees have a higher employment rate.

The refugee immigrants are compared to benchmark groups. The first group consists of randomly selected natives (cohort 1) in the same age groups as the refugees. The second is a matched control group of native-born workers (cohort 2) which was created by coarsened exact matching (CEM) (Iacus, King and Porro, 2012; Blackwell, Iacus, King and Porro, 2009; King, Lucas and Nielsen, 2017). This method creates comparable cohorts of natives and refugees based on values of gender, marital status, education, parenthood, region where the person lives (district) and birth year.[10]

Following Acemoglu and Autor (2011), we classify all workers into four task categories, defined in Table S2 in the online appendix: (1) cognitive non-routine work tasks (professionals, managers and technicians), (2) cognitive routine tasks (office and administrative support and sales), (3) manual non-routine tasks (personal care, personal service, protective service, food and cleaning), and (4) manual routine tasks (production, craft, repair, operators, fabricators and laborers).

Similar to previous Swedish immigration studies (Erikson, Nordström Skans, Sjögren and Åslund, 2007; Åslund, Forslund and Liljeberg, 2017) we study individuals who are established in the Swedish labor market defined by a particular earnings threshold. We define an established worker as having wage earnings above a 60% threshold of the monthly median labor income in the respective year, controlling for gender. This threshold value allows for low-paid full-time jobs and rules out short temporary jobs that otherwise could bias our results. In

[10] Table S1 in the online appendix reports statistics for the coarsened exact matching.

the robustness section, we test for the sensitivity of our threshold definition.

The study considers individuals with employment over a 20-year period, in which the average age of the workers is above 40 years and the work experience is in the range of 9-14 years. Experience is measured as the cumulative number of years with labor income as the main source of income. We observe workers in six different industry classifications, five different firm sizes, six types of municipalities, and five regions. Using information on the highest educational attainment, we classify the individuals into six categories, from primary school to doctoral degree.

2.2 Descriptive findings

Table 2 shows that over the 2003–2013 period, on average 85% of the matched natives were employed, while 72% of the European refugees and 60% of non-European refugees were employed. The employment rate of the pre-1990 refugee cohort is 65%. Table 2 also reveals that about 88% of employed individuals of the matched native cohort are established in the labor market, while the shares for the refugee cohorts are lower with non-European refugees being lowest with about 75%. The share of individuals with Swedish citizenship is lowest for pre-1990 refugees at 92%, while for natives it is more than 99%.

Table 3 reports how workers in population groups are distributed across occupational task groups. Among matched natives, about 49% of workers are occupied with cognitive non-routine tasks. Closest to this share are pre-1990 refugees with a 34% share. The lowest share is observed for European refugees, while individuals from this group are most likely to work with manual routine tasks (42% vs. 24% for the matched natives). Among the non-European refugees, most work with manual non-routine tasks: 38% vs. 15% among matched natives.

Table 4 displays the average normalized wage earnings for the different population groups across occupational task groups, scaled to median wages in each year. There are significant differences for the first occupational task category of cognitive non-routine tasks. While the matched group of natives have wages 57% higher than the median wage in cognitive non-routine occupations, European refugees have only 25% higher wages, while non-European and pre-1990 refugees have 34% and 38% higher wages, respectively. However, for manual non-routine tasks these two groups have higher wages than native-born workers.

Table 5 shows the frequency of occupations with cognitive non-routine tasks for the different population groups. While for natives the most frequent occupation is technical and commercial sales representatives, nursing associate professionals are most frequently observed for European and non-European refugees. For the pre-1990 refugees, medical doctors constitute the largest group within cognitive non-routine occupational tasks.

3 Empirical strategy

The main identification strategy of this paper is based on matching the refugee immigrants with a sample of comparable natives on observable characteristics one year before the observation period for labor market outcomes.

Using the matched sample, we first estimate a multinomial logit (MNL) model to examine the likelihood that a person belongs to a specific occupational task category. The MNL model determines the impact of variables on the probability of observing each of four alternative outcomes of each characteristic. For worker i in group j at time t, the probability of membership in the alternative task category k is conditional on regressors $\mathbf{x}_{it}$, $\mathbf{q}_{it}$ and $\mathbf{z}_{it}$:

$$Pr[y_{i,t} = k] = \Psi(\gamma_0 + \gamma_1 m_{it} + \gamma_2 x_{it} + \gamma_3 q_{it} + \gamma_4 z_{it} \ + \epsilon_{it}), \ \ k\text{=1,}\ldots\text{,4} \tag{1}$$

where γ_1 captures the effects of group (randomly selected natives, matched natives, European refugees, non-European refugees and pre-1990 refugees), while γ_2 denotes effects of individual characteristics, γ_3 the effects of firm characteristics, γ_4 the impacts of regional characteristics, and ϵ_{it} is an idiosyncratic error term.

We then explain the wage earnings differences between individuals for each of the occupational task categories, using the correlated random effects (CRE) approach (Mundlak, 1978; Wooldridge, 2010). This estimation method has the advantage over the fixed effects model in that we can identify the effects of time-invariant variables, such as being a refugee immigrant, on wage earnings. The CRE approach relaxes the restrictive assumptions of the random effects model in that the unobserved heterogeneity term need not be uncorrelated with other explanatory variables, as those correlations are being modeled.

The CRE model can be written as follows (Schunck, 2013; Schunck and Perales, 2017):

$$y_{it} = \beta_0 + \beta_1 m_{it} + \beta_2 x_{it} + \beta_3 c_i + \pi \bar{x}_i + \mu_i + \epsilon_{it} \tag{2}$$

where y_{it} is the normalized monthly wage earnings of person i, β_1 shows the outcomes for the groups of workers, β_2 shows the within effect on the outcome for time-varying controls x, β_3 is the between effect of time-invariant controls c, π expresses the difference between within and between variation for mean values of the controls, μ_i are individual random effects and ϵ_{it} is an idiosyncratic error term. The controls contain individual, firm and regional characteristics.

The wage model is estimated across all occupational tasks, including the occupational task category as a time-varying control variable, yielding both within and between estimates. We then estimate the model separately for each task category.

Based on the estimates of the wage model, we conduct two additional analyses. To determine the contribution of each of the explanatory factors to the explained wage differences in the CRE equation, we apply the Owen–Shapley R^2 decomposition, following Huettner and Sunder (2012). We also use the Blinder—Oaxaca approach (Blackwell, Iacus, King and Porro, 2009; Oaxaca, 1973) to decompose the observed wage difference between matched natives and refugee workers into explained and unexplained parts.[11]

4 Empirical results

In this section, we present the results of our occupational sorting approach to examine the wage gap of refugee workers. The labor market outcomes are observed over the period 2003–2013 while individual characteristics are available starting in 1999. The refugee sample of almost 100,000 individuals is matched with a similar-sized group of native individuals using individual characteris-

[11] The details of these decompositions are outlined in Section II in the online appendix.

tics for the year 2002. We also consider a benchmark group of about 100,000 randomly selected natives to enable a comparison between matched natives and random natives. The observed differences in outcomes between these two groups can be attributed to differences of characteristics. The total sample of worker-years in the regressions is nearly two million.

4.1 Occupational sorting

The first set of results concerns the probability that a person is employed in a particular occupational task category. Table 6 presents the average marginal effects (AMEs) from the multinomial logit estimation specified by Eq. 1.[12] Controlling for individuals' characteristics, firm size, industry and region, we find that refugees are significantly less likely to work with cognitive non-routine tasks than matched natives. Refugees are more likely to work with manual tasks. As expected, workers living in cities or metropolitan regions and those with more experience and education have a higher propensity to be employed in well-paid cognitive non-routine occupations.

Figure 1 in the online appendix displays conditional marginal effects[13] from interactions with year dummies. The probability for refugees to hold a job in one of the cognitive task categories was 15-25% lower compared to matched natives in the beginning of the period. This gap was only moderately reduced after 10 years. In contrast, refugee immigrants are significantly more likely to

[12] This regression does not report results for the benchmark group of natives as they constitute the base category. The group effects are relative to the benchmark group.

[13] Control variables are gender, municipality of work, marital status, number of children, age category, experience, highest education qualification attained, size of work establishment, industry classification and year.

work in occupations involving manual tasks during the entire period.

It is evident from the sorting model that refugees face obstacles entering the higher-paying cognitive task occupations. This can be due to factors not captured by the model such as time-varying unobserved characteristics of individuals, firms, regions, and institutions. Discrimination in the hiring process could also play a role here.

4.2 Wage earnings

Table 7 displays the estimation results from the wage regression based on Eq. 2. For brevity, only the key coefficients are reported. Variables' suffixes (w) and (b) indicate within and between estimates separating time-variant from time-invariant factors. Column 1 presents results of the wage model for all occupational tasks, controlling for task category, and columns 2–5 display the estimates of the wage equation for the different occupational tasks separately.

The previous literature reports wage gaps in the order of 20-40%. However, our approach of using a matched control group of natives and accounting for the overrepresentation of refugee workers in low-paid occupations shows substantially smaller differences in wages.

Remarkably, the results display not only small gaps but also an inverse gap for European refugees. With randomly selected natives as the reference category, the point estimate is 0.036 for European refugees, compared to 0.018 for matched natives. The corresponding estimates for non-European and pre-1990 refugees are -0.021 and -0.041, respectively. Thus, our results show that the conditional wage difference between refugee-workers and matched native workers varies between +2% and -6%.

The results are closest to those of comparable literature for the cognitive non-routine task group, reported in column 2. The earnings of refugees are 3–9% lower than of the native reference group, and 6–12% lower than the matched native control group. European refugees have 2–5% higher earnings than natives in the three other task categories, as displayed in columns 3–5. Column 3 shows that the average earnings level for cognitive routine tasks differs by only 1% between matched natives, non-European refugees, and pre-1990 refugees. The non-European and pre-1990 refugees have 5% higher earnings than matched natives in manual non-routine occupations (column 4) and 2% lower in manual routine task occupations (column 5).

The within and between regression estimates in column 1 predict the short-term and long-term effect of switching from the manual routine category to one of the other task categories. As can be seen, the earnings effects of switching the occupational task is rather small with one exception: the short-term effect of switching to cognitive non-routine tasks is 5% on average. In addition, the between estimates show that the long-term difference is almost 30%. Furthermore, the effect of an additional year of experience is highest for cognitive non-routine tasks and lowest for manual non-routine tasks. The coefficient estimate of 0.49 for masters' level education in column 2 implies that the return on a master's degree is about 40% for cognitive non-routine tasks compared to primary education.

Another notable finding, in the bottom part of the table, is that the between R^2s are much higher than the within R^2s. The difference between the first column and the other columns shows that the occupational task category has considerable explanatory power for wage differences between individuals. In Figure 2 in the online appendix, the effect from a task category is interacted with the year dummies to examine how the difference evolves over time. The first panel in the upper left corner confirms that the largest earnings gap is associated with cognitive non-routine tasks, and this gap is persistent over time. Figure 2 also shows that the difference in wage earnings is small for the other three occupational tasks, with a tendency towards convergence over time.

4.3 Owen–Shapley R-squared decomposition

Using the Owen–Shapley R^2 decomposition approach described in Eq. (S2), we analyze the marginal contribution to the explained variation in the wage-earnings outcome between refugee and natives workers for the entire labor market presented in Table 7. The overall R^2 of our model, reported in Table 8 column 1, is 0.123. Column 2 shows that 29% of the variation in wages between refugee and native workers can be attributed to occupational sorting, 16% to education, 15% to gender and 11% to work experience. "Other controls", including firm size, civil status, place of living and family characteristics, account for 27% of the wage gap, and the five categories of refugee and native cohorts capture the remaining 3%.

4.4 Blinder–Oaxaca decomposition

While we used the Owen–Shapley analysis above to decompose the CRE results reported in Table 7 into their relative contributions to the explained variation of wage earnings, the Blinder–Oaxaca decomposition described by Eq. S3 in the online appendix examines how much of the observed differences in wages between matched natives and refugees can be explained from their observed characteristics. The results are shown in Table 9. The lower part of the table separates estimated wage differences into explained and unexplained parts.

Over all occupations and for all three refugee groups, the first column of Table 9 shows an unconditional wage gap of 22.6% between matched natives and forced migrants. This is in line with findings in the existing literature. Using our wage model we are able to explain almost the entire gap. The unexplained wage differences between refugees and matched natives correspond to 2.1% lower wages for all refugees in Table 9.

Tables 10 and 11 delineate the cognitive non-routine task into twelve subgroups. In three occupations, primary education teaching associate professionals (Table 10, column 3), doctors (Table 10, column 5), and non-specialist nurses (Table 11, column 1) employees with a refugee background have higher unconditional wages compared to native employees. Controlling for individual and firm heterogeneity, the unexplained estimates show that an inverse wage gap remains in these occupations, as we also find higher relative wages for the refugee group in the subgroup computer system designers and analysts (Table 10, column 1). In three other high skilled occupations, nursing associate professionals, computer assistants (Table 10, columns 2 and 4), and mechanical engineering technicians (Table 10, column 4), the conditional wage gap is only 1–2%. Thus, using wages as a proxy for productivity, the tables suggest immigrant workers may be an important labor market resource in several STEM (science, technology, engineering and mathematics) occupations where many companies face difficulties in recruiting skilled personnel.

Tables 10 and 11 reveal that the unexplained part of the wage differences is significantly larger than the explained part for several of the cognitive non-routine subgroups. What does this reflect? These results might be indicative of discrimination in the labor market. However, as the unexplained part is not uniformly negative, other factors may matter. Refugees earn higher wages than predicted by the model in the education and health care sectors, while we find the opposite among technicians, engineers and public administrators. It is possible to trace a public–private sector dimension in this difference between the work tasks which could imply greater discrimination in the private sector. Another tentative explanation considers unobserved abilities related to the impact of ongoing technological change on the demand for labor. Freeman, Ganguli and Handel (2020) find that the within–occupation impact of technological changes dominated changes between occupations in the U.S. economy over the period 2005–2015. If this pattern is also relevant for the Swedish economy, as is likely, workers with greater ability are more prone to switch to new, more productive and higher-paid job tasks within the cognitive non-routine group.

For the three other occupation categories in Table 9, the wage gap between natives and refugees is substantially smaller for cognitive non-routine tasks, with explained wage differences generally larger than the unexplained differences. Manual non-routine occupations make a notable exception. All three refugee categories in the study earn more than natives in these occupations, and the unexplained differences are significantly larger than the explained differences. Similar to our discussion above, unobserved ability might contribute to the results. If this is the case, refugees may have higher abilities compared to natives in non-routine manual job tasks.

5 Robustness checks

Table 8 shows that years of labor market experience is an important determinant of the differences in wage income between natives and refugees in our results. Our first robustness check considers the sensitivity of results to an alternative definition of experience. In the original analysis, we count experience as the number of years an individual has labor income starting in 1993.

Because there may be problems with this measure as it does not capture the intensity of work effort, we imposed a restriction for their establishment on the labor market, defined as wage earnings above a 60% threshold of the monthly median labor income in the respective year.

As a robustness test, we reestimated the Blinder–Oaxaca decompositions without this restriction, defining work experience as the number of years when an individual has reported income. With this definition of work experience, the unexplained wage differences between natives and refugees increased, but the relevance of experience prevailed.

An additional robustness check for the worker's experience variable was to consider only individuals with employment during the period 1998–2013 rather than the period 1993–2013 that we consider in the main analysis. The justification for this test is the large initial difference in the employment rate between refugees and other immigrants. It takes several years for refugees to establish themselves on the labor market. Comparing the result for experience between the 1998–2013 period and the 1993–2013 period shows that the explanatory power for work experience increases when the time window is extended by five years.

Our interpretation of the two sensitivity tests is that labor market experience is a significant factor influencing the relative wages for refugee immigrants. Thus, the fact that refugees are often constrained to enter the labor market in whatever job is available in the beginning of their stay in the host country, has relevance for work experience as a decisive factor for later wage differences.

There are two potential concerns regarding these analyses. The first is that accumulated work experience might be endogenous, affected by unobserved factors such as ability or motivation. Furthermore, it is plausible that accumulated work experience is affected by wage income. To address this concern, we implemented several instrumental variables (IV) approaches. The first instrument we use in these tests is the occurrence of having twin children. Twin children can be found 850 times in our sample.[14] We define two instruments for the tests: having twin children of ages 0–3 years and having twin children of ages 4–6 years. As expected, we find that having twins of ages between 4 and 6 years also reduces work experience, but by a significantly lower extent compared to having twins between 0–3 years old. The Hansen J test of overidentifying restrictions supports the validity of the IVs at conventional levels, and weak instruments tests are satisfactory. [15, 16]

[14] While our dataset does not provide direct information on having twins, we infer their presence indirectly from the change of the number of children with ages 0–3 years. If this change is 2 or more in a year, we classify this is as an indication of having or adopting twin children. Although Sweden allows for generous benefits while being on parental leave, having small children below the general school age of 7 years reduces accumulated work experience, in particular for women. This effect is even stronger with twins, so that having twins exerts a negative shock to work experience. As adding twins to the family is generally a random event, this satisfies the IV exogeneity assumption.

[15] Note that these IV estimates do not indicate endogeneity of experience at any reasonable level of significance.

[16] The IV–GMM estimations have been performed with Stata's `xtivreg2` command and are available from the authors upon request.

We are able to conduct additional endogeneity tests regarding work experience for the refugee cohorts. We utilize the fact that the asylum decision was granted to individuals in different quarters of the year, as the length of processing times vary. There is also a seasonal tendency in the total number of asylum decisions, with fewer decisions in summer and more at the end of the year. Interestingly, the calendar quarter of positive asylum decision also affects the accumulated work experience in later years. Persons that have obtained their asylum decision in the first quarter of the year have on average about half a year more accumulated work experience compared to refugees who obtained their decision in the third or fourth quarters of the year. We base the test of endogeneity of work experience on using the CRE model for wage income and inserting the residuals from the first-stage regression as an additional regressor. By doing so, this regression equation becomes a control function approach. As the coefficient on the residual is not significant, there is no evidence supporting the endogeneity of experience from these tests using the calendar quarter of asylum decision as exogenous variation.

A second potential concern is that we might overestimate the effect of belonging to occupational task group 1. A person in this group might have earned a higher wage than others in other task groups, leading to a selection of persons with higher ability into task group 1. We address this concern by using a model that predicts whether a person is working in task group 1 or not. As an excluded instrument, we use the initial random allocation of refugees to regions, which is the region where an asylum seeker was first registered in Sweden. To reduce their impact on metropolitan areas, arriving refugees were systematically located across smaller cities and rural districts of Sweden. For natives, we use the municipality where a person was registered in 1990. For younger individuals, this could be the municipality where the person is born. We classify the municipalities into the six categories shown in Table 1.

Our probit model results highlight that persons that were initially located in metropolitan or densely populated regions have respectively a 52% and 31% higher probability to work in task group 1 in later years compared to persons initially located in remote regions. The results show that the error terms of the selection equation and the wage outcome equation have a low negative and significant correlation. More importantly, in the full model, the coefficient of belonging to task group 1 on wage income increases from 4.8 to 6.4 percent. This suggests that we most likely do not overestimate, but rather underestimate, the effect of belonging to task group 1 on the wage.[17] However, the difference in point estimates between these models is not statistically significant.

As a further robustness check, we consider the impact of applying the CEM approach vis-à-vis the commonly used propensity score matching method (Caliendo and Kopeinig, 2008). We obtained qualitatively similar results.

6 Conclusions

An aging population and shortages of labor in cognitive as well as manual occupations pose challenges for productivity, innovation, and growth in many OECD countries. Using administrative register data for Sweden and observations at the work-task level, our study reveals that refugee immigrants have

[17] In this case, we use Roodman's `cmp` Stata command (Roodman, 2011) to estimate a probit model that explains the likelihood of a person to work in task group 1 jointly with the wage income equation. The results are available from the authors upon request.

the potential to alleviate this problem. While the overall wage difference between native-born workers and refugee immigrants is substantial, the empirical results show that occupational sorting, as classified by the skill-biased technical change theory, accounts for the largest share of the estimated wage difference between native and refugee immigrant workers. As refugee workers are less likely to be employed in high-paid jobs and more likely to be sorted into low-skilled jobs than comparable native-born workers with similar individual and family characteristics, the average wage of refugees is less than of native workers. But when looking at the different occupations separately, the estimated wage differences are small or non-existent, which show similar productivity levels between native-born and refugee workers. In several STEM occupations, commonly regarded as strategic for innovation-driven economies and facing shortages of skilled labor, we find only marginal wage differences or even higher wages for refugee workers. Using wages as a proxy for productivity, these findings indicate that facilitating the access of refugee immigrants to high-skill jobs could help to alleviate the problems of labor shortages, enhancing productivity, innovation, and growth of OECD economies.

Our work raises several interesting questions for future research that could include both refugee and economic migrants. First, as many firms face shortages of skilled STEM workers, this paper shows that immigrants may be a target for high-tech job recruitment. An interesting question to study is whether a higher share of immigrants already employed by a STEM firm increases the probability that a new employee is an immigrant. A related research question is whether STEM businesses managed by immigrants are more likely to recruit immigrants than firms managed by natives.

References

Acemoglu, D. and Autor, D. (2011), Skills, tasks and technologies: Implications for employment and earnings, *in* D. Card and O. Ashenfelter, eds, 'Handbook of Labor Economics', Vol. 4b, Elsevier, pp. 1043–1171.

Acemoglu, D. and Restrepo, P. (2018), 'The race between man and machine: Implications of technology for growth, factor shares, and employment', *American Economic Review* **108**(6), 1488–1542.

Ager, A. and Strang, A. (2008), 'Understanding integration: A conceptual framework', *Journal of Refugee Studies* **21**(2), 166–191.

Åslund, O., Forslund, A. and Liljeberg, L. (2017), Labour market entry of non-labour migrants–Swedish evidence, *in* B. Bratsberg, O. Raaum, K. Røed and O. Åslund, eds, 'Nordic Economic Policy Review: Labour Market Integration in the Nordic Countries', Nordisk Ministerråd, pp. 115–158.

Auer, D. (2018), 'Language roulette–the effect of random placement on refugees' labour market integration', *Journal of Ethnic and Migration Studies* **44**(3), 341–362.

Autor, D. H., Levy, F. and Murnane, R. J. (2003), 'The skill content of recent technological change: An empirical exploration', *Quarterly Journal of Economics* **118**(4), 1279–1333.

Balkan, B. and Tumen, S. (2016), 'Immigration and prices: quasi-experimental evidence from Syrian refugees in Turkey', *Journal of Population Economics* **29**(3), 657–686.

Bevelander, P. (2020), 'Integrating refugees into labor markets', *IZA World of Labor* .

Blackwell, M., Iacus, S. M., King, G. and Porro, G. (2009), 'CEM: Coarsened exact matching in Stata', *Stata Journal* **9**(4), 524–546.

Blinder, A. S. (1973), 'Wage discrimination: Reduced form and structural estimates', *The Journal of Human Resources* **8**(4), 436–455.
URL: *http://www.jstor.org/stable/144855*

Borjas, G. J. and Monras, J. (2017), 'The labour market consequences of refugee supply shocks', *Economic Policy* **32**(91), 361–413.

Brell, C., Dustmann, C. and Preston, I. (2020), 'The labor market integration of refugee migrants in high-income countries', *Journal of Economic Perspectives* **34**(1), 94–121.

Caliendo, M. and Kopeinig, S. (2008), 'Some practical guidance for the implementation of propensity score matching', *Journal of Economic Surveys* **22**(1), 31–72.
URL: *https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-6419.2007.00527.x*

Card, D. (1990), 'The impact of the Mariel boatlift on the Miami labor market', *ILR Review* **43**(2), 245–257.

Chin, A. and Cortes, K. E. (2015), The refugee/asylum seeker, *in* B. Chiswick and P. Miller, eds, 'Handbook of the economics of international migration', Vol. 1, Elsevier, pp. 585–658.

Clemens, M. A. and Hunt, J. (2019), 'The labor market effects of refugee waves: reconciling conflicting results', *ILR Review* **72**(4), 818–857.

Connor, P. (2010), 'Explaining the refugee gap: Economic outcomes of refugees versus other immigrants', *Journal of Refugee Studies* **23**(3), 377–397.

Cortes, K. E. (2004), 'Are Refugees Different from Economic Immigrants? Some Empirical Evidence on the Heterogeneity of Immigrant Groups in the United States', *Review of Economics and Statistics* **86**(2), 465–480.
URL: *https://doi.org/10.1162/003465304323031058*

De Vroome, T. and Van Tubergen, F. (2010), 'The employment experience of refugees in the Netherlands', *International Migration Review* **44**(2), 376–403.

Dustmann, C. and Görlach, J.-S. (2016), 'The economics of temporary migrations', *Journal of Economic Literature* **54**(1), 98–136.

Erikson, R., Nordström Skans, O., Sjögren, A. and Åslund, O. (2007), 'Ungdomars och invandrades inträde på arbetsmarknaden 1985–2003', *IFAU rapport* **18**.

Evans, W. N. and Fitzgerald, D. (2017), The economic and social outcomes of refugees in the United States: Evidence from the ACS, Technical report, National Bureau of Economic Research Working Paper 23498.

Fasani, F., Frattini, T. and Minale, L. (2018), '(the struggle for) refugee integration into the labour market: Evidence from europe', *Centro Studi Luca d'Agliano Development Studies Working Paper* (435).

Foged, M. and Peri, G. (2016), 'Immigrants' effect on native workers: New analysis on longitudinal data', *American Economic Journal: Applied Economics* **8**(2), 1–34.

Freeman, R. B., Ganguli, I. and Handel, M. J. (2020), 'Within-occupation changes dominate changes in what workers do: A shift-share decomposition, 2005–2015', *AEA Papers and Proceedings* **110**, 394–399.

Griliches, Z. and Mairesse, J. (1995), 'Production functions: the search for identification'.

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

Huettner, F. and Sunder, M. (2012), 'Axiomatic arguments for decomposing goodness of fit according to shapley and owen values', *Electronic Journal of Statistics* **6**, 1239–1250.

Iacus, S. M., King, G. and Porro, G. (2012), 'Causal inference without balance checking: Coarsened exact matching', *Political Analysis* **20**(1), 1–24.

Juarez, F. C. (2012), 'SHAPLEY2: Stata module to compute additive decomposition of estimation statistics by regressors or groups of regressors', Statistical Software Components, Boston College Department of Economics.
URL: *https://ideas.repec.org/c/boc/bocode/s457543.html*

King, G., Lucas, C. and Nielsen, R. A. (2017), 'The Balance-Sample Size Frontier in Matching Methods for Causal Inference', *American Journal of Political Science* **61**(2), 473–489.

Kuznets, S. (1960), Population change and aggregate output, *in* 'Demographic and economic change in developed countries', Columbia University Press, pp. 324–351.

Lundborg, P. (2013), 'Refugees' employment integration in Sweden: Cultural distance and labor market performance', *Review of International Economics* **21**(2), 219–232.

Marbach, M., Hainmueller, J. and Hangartner, D. (2018), 'The long-term impact of employment bans on the economic integration of refugees', *Science Advances* **4**(9), eaap9519.

Mundlak, Y. (1978), 'On the pooling of time series and cross section data', *Econometrica* **46**(1), 69–85.
URL: *http://www.jstor.org/stable/1913646*

Murnane, R. J., Willett, J. B. and Levy, F. (1995), 'The growing importance of cognitive skills in wage determination', *The review of economics and statistics* pp. 251–266.

Oaxaca, R. (1973), 'Male-female wage differentials in urban labor markets', *International Economic Review* **14**(3), 693–709.
URL: *http://www.jstor.org/stable/2525981*

Peri, G. (2016), 'Immigrants, productivity, and labor markets', *Journal of economic perspectives* **30**(4), 3–30.

Peri, G. and Yasenov, V. (2019), 'The labor market effects of a refugee wave: Synthetic control method meets the Mariel boatlift', *Journal of Human Resources* **54**(2), 267–309.

Roodman, D. (2011), 'Fitting fully observed recursive mixed-process models with cmp', *Stata Journal* **11**(2), 159–206.

Ruiz, I. and Vargas-Silva, C. (2018), 'Differences in labour market outcomes between natives, refugees and other migrants in the UK', *Journal of Economic Geography* **18**(4), 855–885.

Schock, K., Böttche, M., Rosner, R., Wenk-Ansohn, M. and Knaevelsrud, C. (2016), 'Impact of new traumatic or stressful life events on pre-existing ptsd in traumatized refugees: Results of a longitudinal study', *European Journal of Psychotraumatology* **7**(1), 32106.

Schunck, R. (2013), 'Within and between estimates in random-effects models: Advantages and drawbacks of correlated random effects and hybrid models', *Stata Journal* **13**(1), 65–76.
URL: *http://www.stata-journal.com/article.html?article=st0283*

Schunck, R. and Perales, F. (2017), 'Within- and between-cluster effects in generalized linear mixed models: A discussion of approaches and the xthybrid command', *Stata Journal* **17**(1), 89–115.

Tumen, S. (2016), 'The economic impact of Syrian refugees on host countries: Quasi-experimental evidence from Turkey', *American Economic Review* **106**(5), 456–60.

Wooldridge, J. M. (2010), *Econometric Analysis of Cross Section and Panel Data*, 2nd edn, MIT Press.

Tables and Figures

Table 1: Variable descriptions

Variable	Definition
population group	**1=native-born**, 2=matched control group of native-born, 3=European refugees, 4=non-European refugees, 5=pre-1990 refugees
occupational task category	1= cognitive non-routine tasks, 2=cognitive routine tasks, 3=manual non-routine tasks, **4=manual routine tasks**
educ	highest educational attainment: **1=primary school**, 2=secondary school, 3=tertiary education (below university degree), 4=bachelor's degree, 5=master's degree, 6=doctoral degree
female	1=women, **0=men**
age	current year minus birth year. In regression models, age is included as categorical variable, 1=age <30, 2=age 30-34, 3=age 35-39, 4=age 40-44, 5=age 45-49, 6=age 50-54, **7=age 55-59**
married	marital status: 1=married, **0=unmarried**
citizenship	Swedish citizenship: 1=yes, **0=no**
kids age 0-3	number of children with age 0-3 years, winsorized at 2, ref category **0 children**
kids age 4-6	number of children with age 4-6 years, winsorized at 2, ref category **0 children**
wage	monthly wage earnings relative to median monthly wage earnings in respective year differentiated by gender
experience	cumulative number of years with labor income as main source of income
ind	1=high-tech manufacturing, 2=medium-tech manufacturing, 3=low-tech manufacturing, 4=high-tech knowledge intensive services (kis), 5=market kis, **6=less knowledge intensive services**
fsize	number of firm's employees, 1=micro$<$1-9, 2=small 10-49, 3=medium 50-249, 4=large 250-999, **5=big$\geq$1000 employees**
muni	settlement type of municipality where a person's workplace is located, 1= metropolitan area/larger city, 2=densely populated, close to larger city, 3=rural region close to larger city, 4=densely populated remote region, 5=rural remotely located region, **6=rural very remotely located region**
region	aggregated from the 21 counties, 1=Stockholm, 2=Scania, 3=Västra Götaland, 4=south, **5=middle and north Sweden**

Notes: Reference category of a categorical variable is shown in **bold**. The data and variables are based on register information retrieved from Statistics Sweden.

Table 2: CEM samples: Employment, labor market establishment, Swedish citizenship, 2003–2013

	natives	matched natives	European refugees	non-European refugees	pre-1990 refugees
fraction employed	0.845	0.843	0.717	0.597	0.650
of which					
fraction established	0.888	0.882	0.870	0.749	0.794
fraction citizens	0.993	0.992	0.993	0.940	0.917
person-year obs	1,079,632	1,079,622	392,528	333,044	320,474

Notes: A person is defined as being established on the labor market if monthly wage earnings $\geq$ 0.6 monthly median wage earnings, differentiated by gender, conditional on being employed. Citizenship indicates being a Swedish citizen.

Table 3: Share of workers from population group j in occupational task category k, 2003–2013

	natives	matched natives	European refugees	non-European refugees	pre-1990 refugees
cognitive non-routine	0.519	0.487	0.201	0.269	0.344
cognitive routine	0.121	0.124	0.091	0.087	0.085
manual non-routine	0.151	0.151	0.287	0.378	0.324
manual routine	0.209	0.238	0.421	0.266	0.247
observations	753,561	735,772	238,621	138,942	153,932

Notes: Only employed persons established on the labor market, see Table 2.

Table 4: Normalized wage earnings for population group j in occupational task category k, 2003–2013

	natives	matched natives	European refugees	non-European refugees	pre-1990 refugees
cognitive non-routine	1.443	1.570	1.250	1.342	1.381
cognitive routine	0.991	1.003	0.960	0.982	1.005
manual non-routine	0.874	0.881	0.865	0.923	0.930
manual routine	1.118	1.122	1.059	1.036	1.079
observations	753,561	735,772	238,621	138,942	153,932

Notes: Wage earnings relative to median wage earnings in respective year. Only established persons, see Table 2.

Table 5: The 10 most frequent occupations by population group within the cognitive non-routine task category (%)

natives		matched natives		European refugees		non-European refugees		pre-1990 refugees	
Technical and commercial sales representatives (3415)	5.37	Technical and commercial sales representatives (3415)	7.03	Nursing associate professionals (2330)	6.19	Nursing associate professionals (2330)	9.83	Medical doctors (2221)	5.57
Primary education teaching associate professionals (3310)	5.33	Computer systems designers and analysts (2131)	4.87	Primary education teaching associate professionals (3310)	4.81	Medical doctors (2221)	8.96	Computer systems designers and analysts (2131)	5.55
Nursing associate professionals (2330)	4.87	Primary education teaching associate professionals (3310)	3.99	Medical doctors (2221)	4.48	Computer systems designers and analysts (2131)	4.70	Nursing associate professionals (2330)	5.51
Computer systems designers and analysts (2131)	4.67	Nursing associate professionals (2330)	3.24	Computer systems designers and analysts (2131)	4.14	Primary education teaching associate professionals (3310)	4.24	Non-specialist nurses (3239)	4.17
Public administration (2470)	2.76	Computer assistants (3121)	2.94	Non-specialist nurses (3239)	3.78	Non-specialist nurses (3239)	3.32	Primary education teaching associate professionals (3310)	4.00
Non-specialist nurses (3239)	2.69	Public administration (2470)	2.33	Public administration (2470)	3.76	Electronics and telecommunications engineers (2144)	2.99	Electronics and telecommunications engineers (2144)	3.12
Administrative secretaries and related associate professionals (3431)	2.43	Physical and engineering science technicians not elsewhere classified (3119)	2.25	Physical and engineering science technicians not elsewhere classified (3119)	3.21	Public administration (2470)	2.94	Computer assistants (3121)	2.94
Computer assistants (3121)	2.42	Administrative secretaries and related associate professionals (3431)	2.06	Mechanical engineering technicians (3115)	3.08	Social service worker (2492)	2.86	Biomedical analytics (3240)	2.63
Medical doctors (2221)	1.92	Mechanical engineering technicians (3115)	1.96	Social service worker (2492)	2.84	Computer assistants (3121)	2.53	Public administration (2470)	2.59
College, university and higher education teaching professionals (2310)	1.81	Directors and chief executives (1210)	1.92	Government social benefits officials (3443)	2.84	General managers in wholesale and retail trade (1314)	2.31	College, university and higher education teaching professionals (2310)	2.20
Cumulative %	34.25		32.58		39.12		44.69		38.30

Notes: Occupation codes using SSYK 96 classification.

Table 6: Marginal effects of being employed in occupational task category k, MNL model

	(1) cogn non-rout	(2) cogn rout	(3) man non-rout	(4) man rout
natives	reference	reference	reference	reference
matched natives	0.011***	-0.003***	-0.002***	-0.006***
	[0.001]	[0.001]	[0.001]	[0.001]
European refugees	-0.168***	-0.026***	0.076***	0.117***
	[0.001]	[0.001]	[0.001]	[0.001]
non-European refugees	-0.195***	-0.029***	0.173***	0.051***
	[0.001]	[0.001]	[0.001]	[0.001]
pre-1990 refugees	-0.125***	-0.033***	0.134***	0.025***
	[0.001]	[0.001]	[0.001]	[0.001]
female	0.014***	0.076***	0.155***	-0.245***
	[0.001]	[0.000]	[0.000]	[0.001]
experience	0.005***	-0.000	-0.005***	-0.000
	[0.000]	[0.000]	[0.000]	[0.000]
secondary school	0.077***	-0.014***	-0.015***	-0.048***
	[0.001]	[0.001]	[0.001]	[0.001]
tertiary school	0.378***	-0.053***	-0.138***	-0.187***
	[0.001]	[0.001]	[0.001]	[0.001]
2 yrs college degree	0.621***	-0.095***	-0.253***	-0.273***
	[0.001]	[0.001]	[0.002]	[0.002]
university degree	0.668***	-0.078***	-0.283***	-0.308***
	[0.001]	[0.001]	[0.002]	[0.002]
fsize micro 1-9	0.009***	0.071***	-0.147***	0.067***
	[0.002]	[0.002]	[0.002]	[0.002]
ind medium-high	0.095***	0.025***	-0.357***	0.238***
	[0.001]	[0.001]	[0.002]	[0.001]
ind medium-low	-0.001	-0.013***	-0.190***	0.204***
	[0.001]	[0.001]	[0.001]	[0.001]
ind low-tech	0.261***	0.103***	-0.238***	-0.125***
	[0.002]	[0.001]	[0.003]	[0.002]
ind KIS	0.144***	-0.014***	-0.062***	-0.069***
	[0.001]	[0.001]	[0.001]	[0.001]
muni metro/city	0.103***	0.032***	-0.066***	-0.070***
	[0.003]	[0.003]	[0.002]	[0.003]
muni dense close city	0.050***	0.015***	-0.044***	-0.021***
	[0.003]	[0.003]	[0.002]	[0.003]
muni rural close city	0.011***	0.004	-0.028***	0.014***

cont.

	(1) cogn non-rout	(2) cogn rout	(3) man non-rout	(4) man rout
	[0.003]	[0.003]	[0.003]	[0.003]
observations	1,936,101			
df (model)	144			
χ^2	1,743,143			
p-value	0.000			

Notes: Standard errors in brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 7: Determinants of wage earnings by occupational category, correlated random effects model

Dep var: *wage*	(1) all occup	(2) cogn non-rout	(3) cogn rout	(4) man non-rout	(5) man rout
time-invariant regressors					
matched native	0.018***	0.026***	-0.004	-0.002	-0.001
	[0.004]	[0.007]	[0.003]	[0.002]	[0.002]
European refug	0.036***	-0.031***	0.021***	0.048***	0.016***
	[0.004]	[0.008]	[0.005]	[0.003]	[0.003]
non-European refug	-0.021***	-0.078***	-0.013*	0.055***	-0.023***
	[0.004]	[0.010]	[0.008]	[0.004]	[0.004]
pre-1990 refug	-0.041***	-0.091***	0.003	0.047***	-0.020***
	[0.005]	[0.009]	[0.008]	[0.004]	[0.004]
female	-0.269***	-0.348***	-0.151***	-0.146***	-0.146***
	[0.004]	[0.006]	[0.003]	[0.003]	[0.003]
time-varying regressors (within estimates)					
non-rout cogn (w)	0.047***				
	[0.002]				
rout cogn (w)	-0.008***				
	[0.003]				
non-rout man (w)	-0.026***				
	[0.003]				
experience (w)	0.070***	0.125***	0.058***	0.031***	0.041***
	[0.002]	[0.005]	[0.004]	[0.002]	[0.003]
experience^2 (w)	-0.001***	-0.001***	-0.001***	-0.000***	-0.000***
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
married (w)	-0.005***	-0.009***	-0.004***	-0.002*	-0.003***
	[0.001]	[0.003]	[0.001]	[0.001]	[0.001]
kid age 0-3: 1 (w)	-0.078***	-0.115***	-0.061***	-0.049***	-0.041***
	[0.002]	[0.003]	[0.003]	[0.002]	[0.002]
kids age 0-3: 2 (w)	-0.140***	-0.201***	-0.102***	-0.090***	-0.070***
	[0.004]	[0.007]	[0.006]	[0.006]	[0.005]
kid age 4-6: 2 (w)	-0.017***	-0.023***	-0.020***	-0.019***	-0.016***
	[0.002]	[0.004]	[0.003]	[0.002]	[0.002]
kids age 4-6: 1 (w)	-0.033***	-0.050***	-0.026***	-0.024***	-0.032***
	[0.005]	[0.009]	[0.010]	[0.005]	[0.005]
educ effects (w)	yes	yes	yes	yes	yes
age effects (w)	yes	yes	yes	yes	yes
firm size effects (w)	yes	yes	yes	yes	yes
industry effects (w)	yes	yes	yes	yes	yes
region effects (w)	yes	yes	yes	yes	yes
citizenship effect (w)	yes	yes	yes	yes	yes

cont.

Dep var: *wage*	(1) all occup	(2) cogn non-rout	(3) cogn rout	(4) man non-rout	(5) man rout
time-varying regressors (between estimates)					
non-rout cogn (b)	0.342***				
	[0.004]				
rout cogn (b)	0.018***				
	[0.004]				
non-rout man (b)	0.035***				
	[0.005]				
experience (b)	-0.027***	-0.048***	-0.019***	-0.001	-0.008***
	[0.002]	[0.004]	[0.003]	[0.002]	[0.002]
experience^2 (b)	0.002***	0.004***	0.002***	0.001***	0.001***
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
age <30 (b)	-0.051***	-0.208***	-0.021	0.029***	0.052***
	[0.014]	[0.029]	[0.019]	[0.010]	[0.012]
age 30-34 (b)	-0.045***	-0.083***	0.022	0.049***	0.054***
	[0.013]	[0.029]	[0.018]	[0.009]	[0.011]
age 35-39 (b)	-0.027**	-0.083***	-0.001	0.037***	0.039***
	[0.013]	[0.028]	[0.018]	[0.009]	[0.010]
age 40-44 (b)	0.052***	0.059*	0.046**	0.050***	0.056***
	[0.014]	[0.030]	[0.018]	[0.008]	[0.010]
age 45-49 (b)	0.047***	0.071**	0.015	0.042***	0.039***
	[0.013]	[0.028]	[0.016]	[0.008]	[0.009]
age 50-54 (b)	0.032*	0.064*	-0.020	0.025**	0.030**
	[0.017]	[0.038]	[0.023]	[0.011]	[0.013]
married (b)	-0.000	-0.007***	-0.001	0.001	-0.001
	[0.001]	[0.003]	[0.002]	[0.001]	[0.001]
educ secondary (b)	0.038***	0.064***	0.023***	0.028***	0.023***
	[0.002]	[0.009]	[0.005]	[0.003]	[0.002]
educ tertiary (b)	0.066***	0.174***	0.075***	0.053***	0.062***
	[0.004]	[0.009]	[0.006]	[0.004]	[0.004]
educ bachelor (b)	0.139***	0.334***	0.135***	0.107***	0.062***
	[0.013]	[0.018]	[0.013]	[0.009]	[0.010]
educ master (b)	0.326***	0.493***	0.210***	0.130***	0.095***
	[0.008]	[0.012]	[0.015]	[0.013]	[0.013]
educ doctoral (b)	0.480***	0.596***	0.177***	0.120***	0.064*
	[0.024]	[0.026]	[0.066]	[0.043]	[0.035]
muni metro/city (b)	0.161***	0.271***	0.116***	0.055***	0.048***
	[0.009]	[0.021]	[0.017]	[0.008]	[0.009]
muni dense close city (b)	0.026***	0.040*	0.041**	0.014*	0.021**
	[0.009]	[0.022]	[0.017]	[0.008]	[0.009]
muni rural close city (b)	0.009	0.015	0.006	-0.003	-0.002

cont.

Dep var: *wage*	(1) all occup	(2) cogn non-rout	(3) cogn rout	(4) man non-rout	(5) man rout
	[0.009]	[0.022]	[0.017]	[0.008]	[0.009]
muni dense remote (b)	0.007	0.008	0.010	0.003	0.013
	[0.009]	[0.022]	[0.018]	[0.009]	[0.009]
muni rural remote (b)	-0.011	-0.018	-0.009	-0.007	-0.018**
	[0.009]	[0.022]	[0.018]	[0.009]	[0.009]
fsize micro 1-9 (b)	-0.142***	-0.216***	0.046***	-0.041***	-0.179***
	[0.011]	[0.019]	[0.010]	[0.008]	[0.010]
fsize small 10-49 (b)	-0.066***	-0.076***	0.096***	-0.035***	-0.107***
	[0.012]	[0.022]	[0.010]	[0.007]	[0.010]
fsize medium 50-249 (b)	-0.076***	-0.110***	0.100***	-0.030***	-0.088***
	[0.011]	[0.018]	[0.010]	[0.007]	[0.010]
fsize large 250-999 (b)	0.008	0.007	0.117***	0.009	-0.022**
	[0.011]	[0.019]	[0.011]	[0.008]	[0.010]
ind high-tech (b)	0.343***	0.398***	0.121***	0.167***	0.020**
	[0.025]	[0.036]	[0.030]	[0.062]	[0.010]
ind medium-high (b)	0.137***	0.200***	0.095***	0.116***	0.010***
	[0.005]	[0.011]	[0.006]	[0.015]	[0.003]
ind medium-low (b)	0.139***	0.295***	0.066***	-0.064***	0.013***
	[0.012]	[0.043]	[0.010]	[0.008]	[0.004]
ind low-tech (b)	0.257***	0.323***	0.017	0.058***	-0.013
	[0.009]	[0.011]	[0.011]	[0.022]	[0.013]
ind KIS (b)	0.268***	0.396***	0.081***	0.020***	-0.024***
	[0.008]	[0.011]	[0.007]	[0.006]	[0.006]
Swedish citizenship	0.037***	0.050**	0.034**	0.041***	0.009
	[0.008]	[0.022]	[0.016]	[0.007]	[0.009]
Constant	1.057***	1.491***	0.827***	0.789***	1.088***
	[0.027]	[0.064]	[0.046]	[0.026]	[0.033]
year effects (b)	yes	yes	yes	yes	yes
kids age 0-3 (b)	yes	yes	yes	yes	yes
kids age 4-6 (b)	yes	yes	yes	yes	yes
observations	1,937,909	852,355	214,165	362,093	483,155
σ_u	0.762	1.098	0.271	0.224	0.364
σ_ϵ	0.744	1.070	0.264	0.218	0.355
ρ	0.346	0.310	0.487	0.450	0.238
individuals	231,828	111,277	36,444	54,354	65,621
df(model)	97	91	91	91	91
R^2 (w)	0.005	0.005	0.013	0.012	0.006
R^2 (b)	0.236	0.172	0.176	0.170	0.140
R^2 (overall)	0.123	0.084	0.110	0.107	0.054

cont.

Dep var: *wage*	(1) all occup	(2) cogn non-rout	(3) cogn rout	(4) man non-rout	(5) man rout

Notes: Cluster-robust standard errors in brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Wage earnings relative to median wage earnings in respective year. (w) indicates within, (b) indicates between. The difference of about 2,000 obs between table 6 and table 7 is due to excluded singletons in Table 7.

Table 8: Owen–Shapley decomposition of overall R^2 in the regression

Control variables	Owen–Shapley	% contribution
Occupational task groups	0.03542	28.8
Educational background	0.01955	15.9
gender	0.01786	14.5
Work experience	0.01366	11.1
Category of workers (5 categories refugee or native)	0.00324	2.6
Other controls (firm size, region, civil status, number of children, etc)	0.03323	27.0
Total	0.12297	100.0

Notes: Notes: Computations are based on column (1) of Table 7. For details, see Huettner and Sunder (2012). We apply the user-written Stata routine by Juarez (2012).

Table 9: Twofold Blinder–Oaxaca wage decomposition for all refugees, 2003–2013

	(1) all occup	(2) cogn non-rout	(3) cogn rout	(4) man non-rout	(5) man rout
matched natives	1.318***	1.596***	1.032***	0.897***	1.146***
	[0.002]	[0.004]	[0.002]	[0.002]	[0.002]
refugees	1.092***	1.345***	1.000***	0.918***	1.082***
	[0.001]	[0.004]	[0.004]	[0.002]	[0.002]
difference	0.226***	0.251***	0.032***	-0.021***	0.064***
	[0.003]	[0.006]	[0.004]	[0.002]	[0.002]
explained	0.205***	0.132***	0.037***	0.022***	0.072***
	[0.004]	[0.008]	[0.004]	[0.003]	[0.003]
unexplained	0.021***	0.119***	-0.005	-0.044***	-0.007*
	[0.005]	[0.009]	[0.006]	[0.004]	[0.004]
N matched natives	706,115	343,808	85,632	102,008	165,575
N refugees	506,922	131,899	43,292	155,749	168,346
Total obs	1,213,037	475,707	128,924	257,757	333,921

Notes: Standard errors in brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Estimations based on correlated random effects model eq. (2). Reference group matched natives. Wage earnings relative to median wage earnings in respective year.

Table 10: Two-fold Blinder-Oaxaca wage decomposition for the most frequent occupations (cognitive non-routine tasks), panel 1

	(1) 21.31	(2) 23.30	(3) 33.10	(4) 31.21	(5) 22.21	(6) 24.70
matched natives	1.733***	1.040***	0.929***	1.472***	2.473***	1.408***
	[0.011]	[0.004]	[0.004]	[0.010]	[0.027]	[0.009]
refugees	1.565***	1.021***	0.939***	1.335***	2.528***	1.191***
	[0.012]	[0.005]	[0.005]	[0.017]	[0.026]	[0.010]
difference	0.168***	0.019***	-0.010	0.137***	-0.055	0.217***
	[0.016]	[0.007]	[0.007]	[0.019]	[0.037]	[0.014]
explained	0.181***	0.012*	0.030***	0.119***	0.249***	0.169***
	[0.022]	[0.007]	[0.008]	[0.020]	[0.038]	[0.017]
unexplained	-0.013	0.006	-0.040***	0.019	-0.304***	0.048**
	[0.026]	[0.009]	[0.010]	[0.026]	[0.051]	[0.021]
N matched natives	15,763	9,814	11,662	9,394	4,562	7,153
N refugees	6,314	7,997	5,274	3,236	8,193	3,569
Total obs	22,077	17,811	16,936	12,630	12,755	10,722

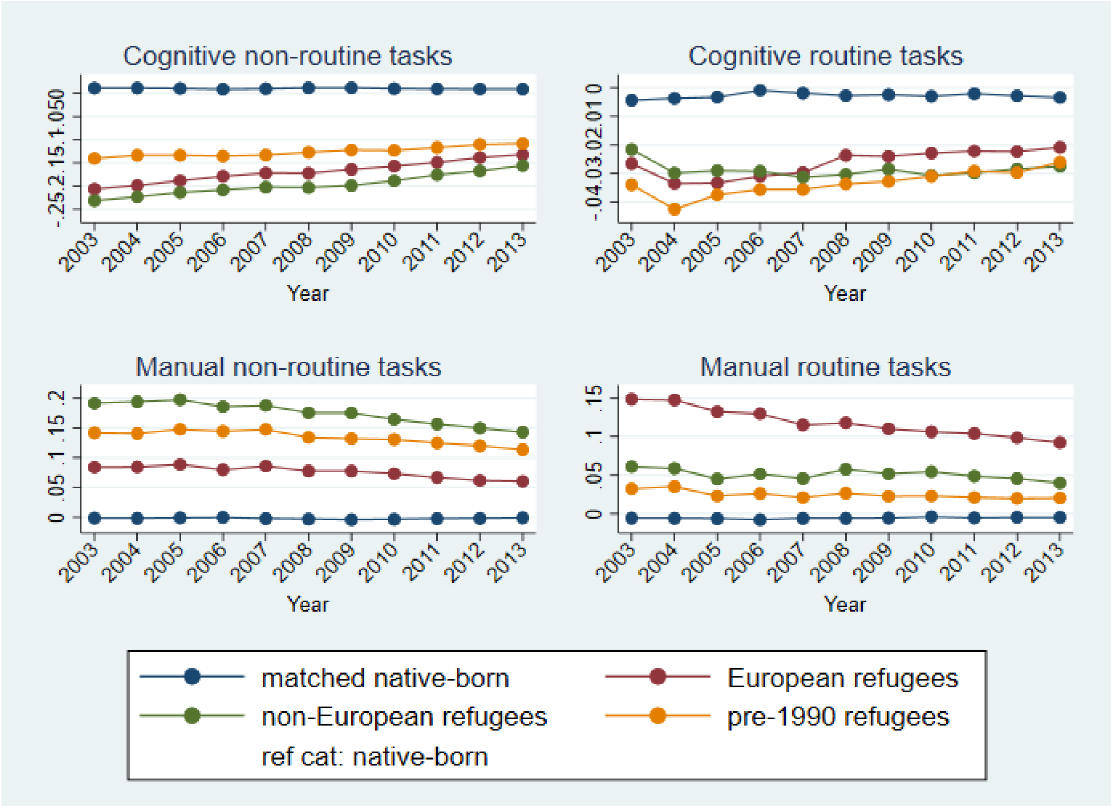
Notes: see Table 9. Occupational codes: 21.31: Computer systems designers and analysts, 23.30: Nursing associate professionals, 33.10: Primary education teaching associate professionals, 31.21: Computer assistants, 22.21: Medical doctors, 24.70: Public administration.

Table 11: Two-fold Blinder-Oaxaca wage decomposition for occupations (cognitive non-routine tasks), panel 2

	(1) 32.39	(2) 31.19	(3) 34.31	(4) 31.15	(5) 21.44	(6) 31.14
matched natives	1.081***	1.484***	1.240***	1.461***	2.031***	1.666***
	[0.006]	[0.010]	[0.012]	[0.012]	[0.031]	[0.016]
refugees	1.178***	1.369***	1.095***	1.414***	1.751***	1.561***
	[0.009]	[0.017]	[0.017]	[0.025]	[0.018]	[0.019]
difference	-0.097***	0.115***	0.145***	0.047*	0.279***	0.105***
	[0.011]	[0.019]	[0.021]	[0.028]	[0.036]	[0.025]
explained	0.027***	-0.009	0.090***	0.029	0.136**	0.030
	[0.009]	[0.022]	[0.023]	[0.031]	[0.064]	[0.042]
unexplained	-0.124***	0.124***	0.055*	0.017	0.143**	0.074
	[0.013]	[0.029]	[0.030]	[0.041]	[0.073]	[0.048]
N matched natives	5,393	6,999	7,044	5,773	3,732	4,978
N refugees	4,816	2,600	2,326	2,462	3,406	2,045
Total obs	10,209	9,599	9,370	8,235	7,138	7,023

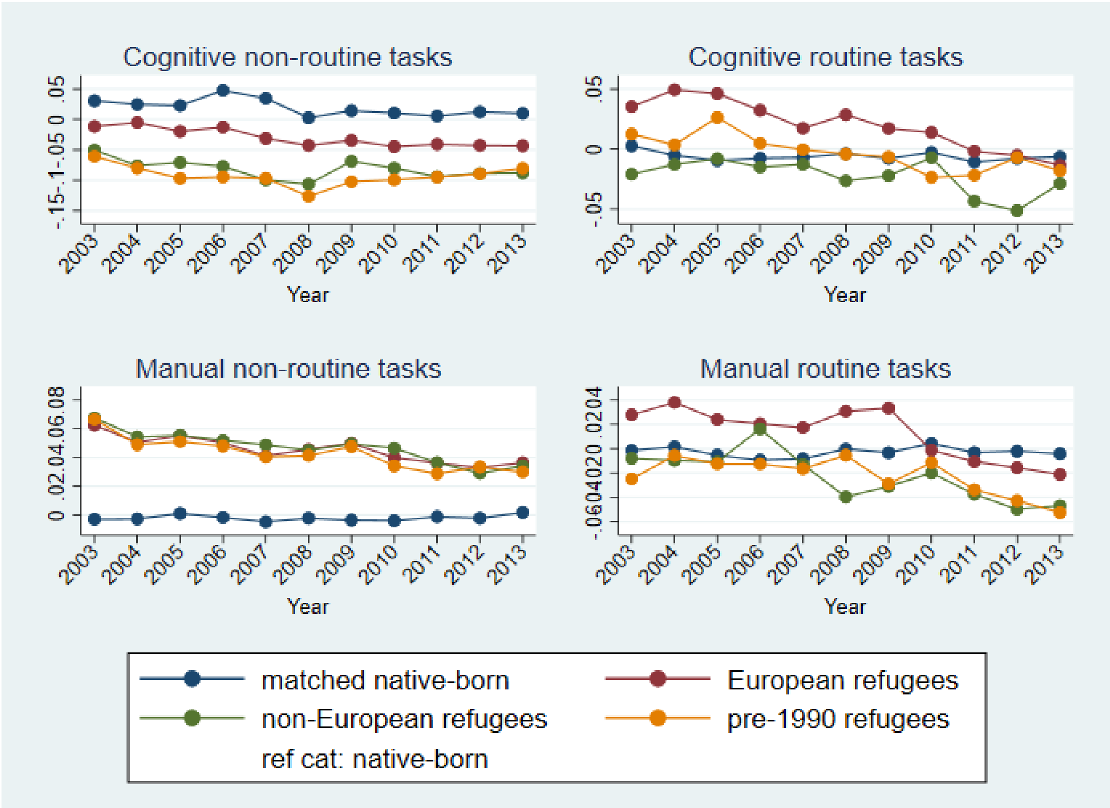
Notes: see Table 9. Occupational codes: 32.39: Non-specialist nurses, 31.19: Physical and engineering science technicians not elsewhere classified, 34.31: Administrative secretaries and related associate professionals, 31.15: Mechanical engineering technicians, 21.44: Electronics and telecommunications engineers, 31.14 Electronics and telecommunications technicians.

Figure 1: Marginal effect of population group on the probability to belong to occupational category k



Notes: Marginal effects from a multinomial logit model with the following control variables: year, gender, municipality of work, marital status, number of children, age category, experience, highest education qualification attained, size of work establishment, industry classification.

Figure 2: Marginal effect of population group on wage earnings in occupational category k



Notes: Marginal effects from a multinomial logit model with the following control variables: year, gender, municipality of work, marital status, number of children, age category, experience, highest education qualification attained, Swedish citizenship, size of work establishment, industry classification.