Labor Supply Shocks and Capital Accumulation:

The Short and Long Run Effects of the Refugee Crisis in Europe

Online Appendix[*]

Lorenzo Caliendo   Luca David Opromolla   Fernando Parro   Alessandro Sforza

January 13, 2023

Data Sources

Survey of Intentions and Perspectives of Refugees from Ukraine

We track the number of refugees from Ukraine in 23 European countries and in a constructed rest of the world using the second round of the survey of intentions and perspectives of refugees from Ukraine UNHCR (2022a), carried out by the United Nations High Commissioner for Refugees (UN-HCR).[1] The data collection exercise, phone-based, web-based or using face-to-face interviews, was carried out between August and September 2022. It includes over 4,800 surveys, each of which used a harmonized questionnaire, which collected data on refugees' demographic profile, including about their place of origin, conditions in current host country and detailed intentions information. Weights have been applied based on the most recent figures available of the number of individual refugees recorded in each country.[2] We focus on refugees that are between 18 and 59 years old; we assign them to a low skill—primary, secondary or no education—or high skill—bachelor, master, doctorate or technical or vocational studies—education groups, and to a group of employed—employed or self-employed/own business— or not employed—apprentice/intern/volunteer, family responsibilities, professional training, retired, student, supporting family business, unemployed, or other. As for the other datasets that are described below, we focus on 23 European countries, which cover about 85 percent of the overall stock of refugees, and a constructed RoW.

The survey data is used to measure the stock of refugees from Ukraine by host country, education level, and employment status.

European Union Labour Force Survey

List of Countries

The sample includes 23 countries and a constructed rest of the world (RoW). The list of countries includes Austria, Belgium, Bulgaria, Cyprus, Czech Republic, Germany, Denmark, Estonia, Spain, France, Greece, Hungary, Italy, Lithuania, Luxembourg, Latvia, The Netherlands, Poland, Portugal, Romania, Slovenia, Slovakia and the U.K., which all were members of the European Union in 2019. Overall, these 23 countries cover about 95 percent of the population of the 28 members of the European Union in 2019.

We assign Finland, Ireland, and Sweden to the RoW aggregate mainly because their EU-LFS country surveys do not contain sufficient information regarding the country of residence and/or the employment status 12 months before the worker was interviewed. Moreover, we assign to the RoW the following countries: Croatia, Iceland, Malta, Norway, and Switzerland. Croatia is not included in the WIOD socio-economic accounts data. Norway, Iceland and Switzerland are not EU28 members. On top of that, the country of origin is always missing for Norway, and half of the times missing for Iceland. Malta presents lots of zeros in the stocks and flows of migrants.

Construction of the Gross Migration Flows

Data on gross migration flows by country of origin, destination, nationality, skill, labor force status, and year are constructed from the micro data of the European Labour Force Survey (EU-LFS) Version 1 released in November 2021 and containing all data up to 2020. For each individual surveyed, the questionnaire reports the country in which the individual resided 12 months before—besides reporting the current country of residence, the year and week in which the individual was interviewed, and a sampling weight that makes the survey representative at the national level. We refer to the country in which the survey was carried out as "destination", and to the country in which the interviewed individual was living 12 months before as "origin". The questionnaire also reports information regarding the age, education, labor force status and nationality of the household. We focus on individuals between 20 and 59 years old, and use the information reported to infer if the individual is a migrant—in case the country where she resides today is different from the one she was residing one year before—as well as the origin country, and the year of migration.

We import the raw micro data into Stata using the Setup_EULFS_1983-2020_y.do and Labels_EULFS_1983-2020.do Stata dofiles made available by the German Microdata Lab of GESIS - Leibniz Institute for the Social Sciences (http://www.gesis.org/en/gml/). We restrict the imported dataset to the following set of variables:

country - Current country of residence (ISO country classification)

countr1y - Country of residence one year before survey (ISO country classification)

refyear - Year of survey

coeff - Yearly weighting factor (Numbers in thousands)

age - Age of interviewed person (in 5-year age bands, 0-4, 5-9,...)

national - Nationality

hat11lev - Highest educational attainment level (ISCED 2011 classification)

ilostat - ILO Work status

wstat1y - Situation with regard to activity one year before survey

Nationality, Labor Force Status, and Education

The EU-LFS contains information on the nationality of the interviewed individuals. However, mainly because of country-specific privacy regulations, the variable "nationality" is limited to 15 country groups. We aggregate all nationalities to two categories: EU28 and Other nationals.

We measure the labor force status in the destination (i.e. in the survey country) using the variable *ilostat*, which measures the labor force status following the guidelines of the International Labour Organization (ILO). We consider two labor force status: employment and non-employment, the latter pooling together unemployed, inactive, and compulsory military service. The labor force status in the origin country (i.e. one year before the survey) is measured through the variable *wstat1y*, which measures the situation with regard to activity one year before survey. Here as well, we consider two aggregate labor force status: employed and not employed. Employed corresponds to "Carries out a job or profession, including unpaid work for a family business or holding, including an apprenticeship or paid traineeship, etc," while not employed pools together unemployed, and a number of inactive statuses including "Pupil, student, further training, unpaid work experience", "In retirement or early retirement or has given up business", "Permanently disabled", "In compulsory military service", "Fulfilling domestic tasks", and "Other inactive person".[3]

The EU-LFS contains information on the education level of the interviewed individuals. Each individual is assigned an education level according to the International Standard Classification of Education (ISCED). The raw data follows ISCED 2011 from the survey year 2014. We use the ISCED classification to split individuals into two education levels, defining as high-skilled all the individuals with at least some tertiary education (ISCED 5). We assign to the low-skilled group the residual workers with education up to post secondary non-tertiary education (below ISCED 1 to ISCED 4). The main reason why we separate categories this way in order to obtain skill groups that are comparable across countries and over time.

Stocks and Flows

Our goal is to measure the stock of people in $t-1$ in each country, by nationality, education level and employment status, and to measure the flows from each country i and employment status v to each country j and employment status u, by nationality and education level.

---

[3]An alternative way to measure the labor force status in the survey country is to use the variable *mainstat*, which is constructed in a way that is more similar to the *wstat1y* variable, but it is missing for some countries. We discuss this in more detail in Caliendo et al. (2021).

The EU-LFS allows us to measure the stock of people in $t-1$ in each country (except the RoW) in two ways. Let's consider a given nationality-skill pair. For each country-labor force status-year cell (i, v, t) we potentially have two separate measures of the stock of individuals: the first measure comes directly from the EU-LFS (i, t) survey; the second measure can be constructed from the set of EU-LFS $\{(i, t+1)\}_i$ surveys for the following year. For example, the Polish survey of 2018 provides a measure of the number of low-skill NMS nationals with labor force status v in Poland in 2018. However, a second measure can be constructed using the surveys for all countries in 2019—including the survey for Poland—reporting immigrants with labor force status v in Poland the year before. Let's define the first measure as $S_{15}^{PL,v}$ and the second measure as $\tilde{S}_{15}^{PL,v}$. If $S_{15}^{PL,v} > \tilde{S}_{15}^{PL,v}$ we can conjecture that the difference $\left(S_{15}^{PL,v} - \tilde{S}_{15}^{PL,v}\right)$ captures migrants from Poland to the RoW (for which there is no survey available). To the contrary, if $S_{15}^{PL,v} < \tilde{S}_{15}^{PL,v}$ we can replace $S_{15}^{PL,v}$ with $\tilde{S}_{15}^{PL,v}$.[4]

Since the EU-LFS is only available for a selected number of countries, we do not observe: (i) flows to the RoW, and (ii) the stock of people in the RoW in $t-1$. We infer flows from any sample country i to the RoW by computing the difference between the stock of people in i and the observed flows from i to any sample country. Provided the difference is positive. If it's negative we update the stock information in $t-1$.

In order to perform the algorithm described above, we need information on the stock of people in the RoW, as well as their distribution across nationalities and labor force status. We use information on population levels and on the share of population between 15 and 64 years old from the World Bank World Development Indicators database to construct the stock of people in the rest of world in 2018.[5] We further use the average year-nationality-skill-labor force status shares from our 23 countries and apply them to the RoW population to split people in the relevant groups for our analysis. Specifically, we apply the average year-nationality-skill-labor force status shares from our 23 countries to a share of the population of the RoW that equals the aggregate population of our 23 countries. The remaining part of the population of the RoW is assigned to the "Other" nationality, with a split between employed and not employed that reflects that average share of employed among "Other" nationals in the 23 EU countries.

Some destination-origin-nationality-skill-year flows have missing values that we interpret as measurement errors in the surveys. We perform the following procedure to, essentially, replace them with tiny positive values. We set the migration shares that are zero to be equal to the minimum positive migration share in the data, conditional on all the relevant information that affects migration decisions in our model. In particular, in our baseline specification we compute the minimum migration share conditional on nationality, skill, labor force status in $t-1$, labor force status in t and year, exploiting the variation across origin and destination countries.

---

[4]In Caliendo et al. (2021) we extend this idea to create an internally consistent dataset of stocks and flows for multiple years.

[5]Total population is based on the *de facto* definition of population, which counts all residents regardless of legal status or citizenship. The values used are midyear estimates.

Trade and Production Data

Bilateral Trade

The bilateral trade flows between each country in the sample are computed using information from the WIOD database (Timmer et al. 2015) for the latest available year 2014. We keep the set of countries consistent with the migration data and we pool all the remaining countries in the rest of the world. Values are in US dollars at current prices.

Share of Labor Compensation in Value Added

We compute the share of labor compensation in value added by skill at the national level using data on value added and labor payment from the Socio Economic Accounts in the WIOD database and the EUKLEMS database.

Nominal Capital Stock

We construct the capital stock for the 23 countries plus the RoW using data from the IMF Investment and Capital Stock Dataset (1960-2019). To construct a measure of total capital stock for a country, we sum three different indicators: the general government capital stock (kgov_n), the private capital stock (kpriv_n), and the public-private partnership (PPP) capital stock (kppp_n). The three variables are in billions of nominal national currency, and we use the exchange rates and deflators from the Penn World Table 10.0 to transform all the values in million US dollars at constant prices. In doing so, we drop some countries for which we do not have information on the exchange rate from the Penn World Tables. Our set of countries represents 99.67 of World GDP in 2014. We use the stock of capital compute the real rental rate of capital, taking the average rental rates over the period 2014-2017.

References

CALIENDO, L., L. D. OPROMOLLA, F. PARRO, AND A. SFORZA (2021): "Goods and factor market integration: a quantitative assessment of the EU enlargement," *Journal of Political Economy*, 129, 3491–3545.

TIMMER, M. P., E. DIETZENBACHER, B. LOS, R. STEHRER, AND G. J. DE VRIES (2015): "An Illustrated User Guide to the World Input–Output Database: the Case of Global Automotive Production," *Review of International Economics*, 23, 575–605.

UNHCR (2022a): "Belgium, Bulgaria, Czech Republic, Germany, France, Italy, Lithuania, Moldova, Poland, Romania, Slovak Republic, Spain, Turkey: Survey of intentions and perspectives of refugees from Ukraine version 2 Sept 2022," .

——— (2022b): "Lives on Hold: Intentions and Perspectives of Refugees from Ukraine," .