Firms Left Behind: Emigration and Firm Productivity[a]

Yvonne Giesing[b], Nadzeya Laurentsyeva[c]

April 15, 2023

Abstract

This paper estimates causal impacts of emigration on firm productivity. We compiled a new industry-level migration dataset and construct an innovative instrument for emigration by exploiting variation in the European labor mobility policies between 2004 and 2016. Using a large firm-level panel from eleven Central and Eastern European countries, we demonstrate that emigration reduces the productivity of an average firm. Additional evidence points to increased skill shortages and the substitution of emigrants with less productive workers. At industry level, however, the exit of the least productive and the entry of more productive firms attenuate the negative effects of emigration.

JEL classification: O15, D24, F22, J24

Keywords: Migration, Firm Productivity, Human Capital, EU Enlargement

[a]We thank Ran Abramitzky, Carsten Eckel, Florian Englmaier, Lisandra Flach, Guido Friebel, Clément Imbert, Takeshi Murooka, Panu Poutvaara, Esteban Rossi-Hansberg, Yona Rubenstein, Julia Schmieder, Monika Schnitzer, Uta Schönberg, Davide Suverato, David Zuchowski, and participants of COPE, GEP/CEPR Postgraduate Conference, Econometric Society European Winter Meeting, Spring Meeting of Young Economists, ifo CEMIR conference, IZA/World Bank conference, IZA Annual Migration Meeting, the annual conferences of EEA, EALE, and German Economic Association, CESifo Area Conference on Employment and Social Protection, World Congress of the Econometric Society, CESifo Venice Summer Institute, and several other seminars at the EBRD, European Commission, University of Munich, Bamberg, Prag and Warwick for helpful comments and suggestions. We thank Julia Gorochovskij for excellent research assistance. Financial support from the DFG Research Training Group 1928 and the CRC TRR 190 is gratefully acknowledged. We also thank the LMU-ifo Economics & Business Data Center for providing access to the Amadeus and Orbis database and various people at national statistical offices for providing immigration data.

[b]Contact: giesing@ifo.de, ifo Institute, University of Munich and CESifo.

[c]Contact: Nadzeya.Laurentsyeva@econ.lmu.de, University of Munich and CESifo.

1 Introduction

"In the Lithuanian town of Panevezys, a shiny new factory [...] sits alone in the local free economic zone. The factory is unable to fill 40 of its jobs, an eighth of the total. That is not because workers in Panevezys are too picky, but because there are fewer and fewer of them."

The Economist (January 19, 2017)

The emigration of skilled workers poses a challenge for many countries, not just the developing world. As workers leave their countries of origin to exploit opportunities abroad, policymakers and firm managers raise concerns about impeding skill shortages and brain drain. These concerns could translate into policies that discourage emigration, create barriers to cross-border labor mobility and hinder regional integration. However, the causal link between skilled emigration and deteriorating firm performance has not been established. The causality could go the other way around with people leaving due to negative economic shocks in their country of origin, or other unobserved factors could trigger lower firm performance and higher emigration. The scarcity of high-quality emigration and firm-level data has constrained the empirical analysis to date. Yet, understanding how firms respond to the outflow of workers is vital for estimating and understanding the economic consequences of emigration and for designing appropriate policies in countries of origin.

This paper investigates the causal effects of emigration on firm performance. We exploit the unique institutional setting of the EU enlargements in 2004, 2007 and 2013 for our identification strategy and analyze the performance of manufacturing firms in eleven Central and Eastern European countries. Following their accession to the EU, these countries experienced unprecedented emigration rates, which induced heated debates about brain drain and its negative economic consequences (Figure A.4). Our main outcome of interest is firm productivity, measured by three indicators: labor productivity, wage-adjusted labor productivity and total factor productivity (TFP). TFP, in particular, has been considered a strong predictor of firms' survival and growth (Bloom, Sadun and Van Reenen, 2012; Bartelsman, Haltiwanger and Scarpetta, 2013). We show that firms in industries exposed to higher emigration experience a drop in all three productivity measures. According

to our main specification, a one percentage point increase in the labor emigration rate leads to a contemporaneous 1.1-1.4 percent decrease in an average firm's labor productivity and a 2.3 percent decrease in TFP. The negative effects persist in the long term. When estimating the emigration effects for a sub-sample of firms over a ten-year period (2003-2013), we find that a one percentage point increase in the emigration rate is associated with a 1.6-3.7 percent lower labor productivity and a 4.6 percent lower TFP for an average firm. Our baseline results are robust to excluding very small or large firms[1], controlling for changes in the sample composition, restricting the sample to certain periods, and using alternative empirical specifications.

The negative productivity effects we find are economically significant. The median increase in the emigration rate over 10 years (3.6 percentage points) translates into a 5.6-12.4 percent drop in a firm's labor productivity and a 15 percent lower TFP in our sample. To compare, Gorodnichenko and Schnitzer (2013) analyze the effect of financial constraints on firm productivity in Central and Eastern European countries and find that a one standard deviation increase in the severity of financial constraints leads to a 21 percent drop in TFP.

To identify causal effects, we construct a new instrument based on variation in the European labor mobility regulations, which were introduced after the 2004, 2007 and 2013 EU eastern enlargements. For a period of seven years after the official accession of new Member States (NMS), the old EU Member States could apply transitional provisions and restrict access to their labor markets for NMS workers. As a result, from 2004 to 2011 (for 2004 accession countries), 2007 to 2014 (for 2007 accession countries), 2013 to 2020 (for the 2013 accession country), labor mobility opportunities for NMS workers within the EU varied depending on their *country of origin*, *destination*, and the *industry* they worked in. We complement these time-varying labor mobility regulations with static data on the *pre-2004* potential to hire migrant workers in the destination industries and proximity between destination and origin countries to construct a free labor mobility variable (FLM) that serves as an instrument for emigration. The instrument is plausibly exogenous to firm productivity in NMS because the detailed industry-, year- and destination-

[1] i.e. excluding firms with less than five or more than 250 employees.

specific regulatory changes were unlikely to be influenced by individual firms in countries of origin. These openings were determined by labor market considerations of the destination countries and were not directed at individual countries of origin but at all accession countries of a given accession year. They were uncorrelated with other immediate accession policy changes, such as the free movement of goods or capital. Furthermore, the level of variation in the instrument allows us to conservatively control for origin- and industry-specific time-varying shocks. Using a new, self-compiled industry-level migration dataset, we show that our instrument is relevant and that it significantly affects labor emigration from NMS.

Firm-level panel data allow us to account for firm heterogeneity and empirically explore the link between firms' characteristics and their sensitivity and adjustment to the emigration of workers. When grouping the incumbent firms into deciles according to their average TFP before 2004 and estimating the emigration effects separately for each productivity decile, we find that the negative effects of emigration on firm performance weaken as we move along the productivity distribution from the least to the most productive firms. We complement firm-level results with the analysis at the industry level, which accounts for the entry and exit of firms. In contrast to the within-firm analysis, we do not detect negative effects of emigration on industry-level labor productivity measures. This is because the adjustment to emigration also takes place at the extensive margin, resulting in the exit of the weakest firms and the entry of new, more productive firms. In the longer term, we document stronger labor reallocation toward more productive firms as firm-size-weighted productivity increases in industries with higher exposure to emigration.

We provide evidence pointing to the mechanisms that can explain our results. Using the EU Commission Business Survey data, we demonstrate that firms in industries that experienced higher workforce emigration are more likely to report skill shortages. Concurrently, we do not observe statistically significant changes in the number of employees, nor in firm personnel costs per employee. While this result might be surprising at first sight, it is consistent with firms replacing emigrated skilled workers with new hires who are less productive. The lower quality of replacement workers (e.g. due to poorer professional qualifications or shorter working experience) is reflected in the lower average wage, which countervails

the upward pressure on firm personnel costs as labor becomes scarcer and hiring costs increase. It is also plausible that new hires are not direct substitutes for those who left due to lower positive externalities (such as firm-specific human capital or spillovers on other workers) that are not reflected in wages but are important for TFP. Additionally, we observe a significant decrease in firms' capital-labor ratio suggesting complementarity between skilled labor and capital investments. While we cannot observe it directly with our data, we can hypothesize that firms' inability to replace emigrated workers with equally skilled ones might translate to lower investments not only in capital but also in new technologies, such as management practices or new information systems, which would be consistent with the persistent negative TFP effects that we find.

This paper relates to several strands of literature. Most importantly, we contribute to the literature studying the effects of emigration on countries of origin by providing firm-level evidence. The brain drain literature has traditionally focused on the negative effects on countries of origin through a loss of human capital (Kapur and McHale, 2005; Docquier, Lohest and Marfouk, 2007). However, further literature has also emphasized potential positive effects arising from higher incentives to invest in tertiary education, as well as from incoming remittances and human capital gains of return migrants (Docquier and Rapoport, 2008; Bollard et al., 2011).[2] Other studies on the economic effects of emigration and brain drain have focused on wages (Docquier, Ozden and Peri, 2014; Dustmann, Frattini and Rosso, 2015; Hafner, 2021), aggregated economic performance (Freeman, 2006; Docquier and Rapoport, 2012; Grossmann and Stadelmann, 2011, 2013; Clemens, 2013), and firm entry and entrepreneurship (Anelli et al., 2022).[3] We contribute with innovative firm-level and industry-level analyses that cover eleven countries of origin and provide evidence on firm heterogeneity and potential mechanisms at play.

While there is a growing migration literature that focuses on the firm as the unit of analysis until now it has investigated the impact of *immigration*.[4] Ottaviano and

[2]Highly skilled emigrants in particular may send large amounts of remittances and thus spur growth (Hunt, 2011, 2015).

[3]Bahar et al. (2022) study the effects of return migration on export behavior. However, in our setting, return migration remains at a low level (Atoyan et al., 2016).

[4]To our knowledge, the only exception is Dicarlo (2022) who studies the effects of emigration to

Peri (2013); Kerr, Kerr and Lincoln (2014); Kerr (2013) encourage the firm-level approach, indicating that it allows us to identify firm adjustment mechanisms and address firm heterogeneity. Both are important for our understanding of how the economic effects of migration are being shaped. Peri (2012); Kerr and Kerr (2013); Kerr, Kerr and Lincoln (2014) study the effects of immigration on firm-level outcomes in the US, and Paserman (2013); Mitaritonna, Orefice and Peri (2017); Ottaviano, Peri and Wright (2018) in France, the UK, and Israel respectively. They find that an increased supply of foreign-born workers positively affects firm productivity due to skill complementarities, faster accumulation of capital, and the specialization of natives in more complex tasks.[5] Our research complements this literature by focusing on the effects of *emigration*.

We contribute to the literature on the determinants of firm productivity (Bartelsman, Haltiwanger and Scarpetta, 2013; Bloom and Van Reenen, 2007; Bloom, Sadun and Van Reenen, 2012; Fox and Smeets, 2011). While firm TFP is often treated as one of the core indicators of economic performance, it is also referred to as "the measure of our ignorance" (Syverson, 2011). We provide empirical evidence for the channel that links firm TFP to skill shortages due to emigration.

We are the first to exploit industry-level variation in the EU labor mobility regulations to causally evaluate the consequences of emigration. Moreover, we assembled a novel industry-specific migration dataset covering EU and EFTA Member States. Ortega and Peri (2013); Grogger and Hanson (2011); Beine, Bourgeon and Bricongne (2019) analyze how mobility regulations have affected migration patterns. Rojas-Romagosa and Bollen (2018) show that the general introduction of the free movement of people in the EU increased migration from new to old Member States. Dustmann, Schönberg and Stuhler (2017) and Beerli et al. (2021) show that granting labor market access for cross-border workers in Germany and Switzerland, respectively, increased cross-border workers' employment, while also affecting the native population's employment and wages as well as firm performance.

In addition, we supplement the literature on the consequences of free labor mo-

Switzerland on firms in Italy.

[5]Further firm-level studies include Dustmann and Glitz (2015) for Germany, Peri, Shih and Sparber (2015) for the US, and Imbert et al. (2022) for China.

bility in the context of the EU enlargement. Following their accession to the EU, Central and Eastern European countries have experienced particularly high emigration: in 2003, the number of Central and Eastern Europeans residing in other EU countries amounted to 2.5 million; by 2018 this number reached nine million (OECD Migration Database, Figure 1). Mayr and Peri (2009) develop a model to study the consequences of European free labor mobility on human capital in the countries of origin. They differentiate between brain drain and brain gain due to return migration and increased incentives to invest in education. Dustmann, Frattini and Rosso (2015) and Elsner (2013) estimate the effects of the post-enlargement emigration on wages in Poland and Lithuania, and find that wages increase for stayers. Caliendo et al. (2021) jointly study the economic effects of trade and labor market integration in the EU, and argue that NMS are the main winners of the EU enlargements. However, there are growing concerns that the emigration of skilled workers has posed severe challenges to countries of origin (Kahanec, 2013; OECD, 2013; Zaiceva, 2014). We supplement this literature by providing nuanced evidence at the firm and industry level.

The paper is organized as follows. The next section presents our conceptual framework. Section 3 describes the institutional setting and the data, followed by Section 4 that presents the empirical specification. Section 5 discusses the results including heterogeneous effects. Section 6 sheds light on the mechanisms and Section 7 provides robustness checks. Section 8 concludes and outlines policy implications.

2 Conceptual Framework

Emigration can affect firm productivity through different channels. This section provides an overview of the links between emigration and firms' labor productivity, wage-adjusted labor productivity and TFP. While the focus of this paper is empirical, this section provides theoretical considerations to navigate through the empirical analysis and to identify potential mechanisms.

As opportunities for emigration among the workforce increase, firms are confronted with stronger competition on the labor market. Consequently, wages rise in

the skill group of those workers likely to emigrate (Dustmann, Frattini and Rosso, 2015). Firms can respond by substituting emigrants with workers from a different skill group, and thus by changing the within-firm skill intensity or by substituting labor with capital. If firms substitute emigrated workers with capital, labor productivity (measured as value-added per employee) should not decline, as labor becomes relatively more productive. Wage-adjusted labor productivity (measured as value-added over personnel costs) will behave similarly but will also reflect the consequences of increased wage levels due to more competition for scarce labor. If emigrants are positively selected and firms substitute them with less skilled workers, both labor and wage-adjusted labor productivity will fall. The effect on the wage-adjusted labor productivity might be attenuated as wages also reflect the quality of employees: the personnel costs are likely to decrease when a firm substitutes a skilled worker with a less skilled worker.

These substitution effects, however, will not be reflected in firm TFP (obtained by dividing output by the weighted average of labor and capital input). There are several other channels through which emigration can influence all three productivity measures. For instance, greater competition in the factor market might encourage firms to adopt better managerial practices to use scarce labor more efficiently. In this case, firm TFP will persistently increase (Bloom and Van Reenen, 2007). Technology adoption will also positively influence labor and wage-adjusted labor productivity. In contrast, if mainly skilled workers emigrate, firms might reduce their investments in new technologies, if they complement skills. This will have persistent negative effects on TFP, as well as on the other two productivity measures. All three productivity measures will also be affected when workers generate positive externalities (such as spillovers on co-workers or accumulated firm-specific human capital), which are not reflected in their wages. Firm productivity measures will be negatively affected when such workers leave. Even if firms manage to hire new qualified workers, these might not be direct substitutes, at least in the short-term, due to the lack of firm-specific skills. This mechanism is in line with Jäger (2016) who observes a lack of substitutability between incumbent skilled workers and skilled workers outside the firm.

When examining productivity measures aggregated at the sector level, we need

to consider additional composition effects due to firms' exit and entry. Emigration might induce the exit of less productive firms, thus resulting in a positive selection of survivors. Emigration can also increase the productivity threshold for market entrants thus leading to fewer entries of new firms (Anelli et al., 2022) and an increase in their initial productivity.

3 Institutional Setting and Data

3.1 Labor Mobility Regulations in the EU

This subsection shows how the gradual opening of EU labor markets created time, country, and industry-level variation in the emigration of NMS citizens. In 2004, ten Central, Eastern and Southern European countries joined the EU: Cyprus, Czech Republic, Estonia, Hungary, Latvia, Lithuania, Malta, Poland, Slovakia and Slovenia. While the free movement of goods and capital was introduced by all countries either prior to or at the time of accession, free labor mobility was initially restricted by certain destination countries. Most EU15[6] countries feared an inflow of cheaper labor. Due to political considerations, the old Member States were therefore allowed to unilaterally limit access to their labor markets for up to seven years. These transitional provisions were applied to all NMS in the same way, except Malta and Cyprus. In 2007, Bulgaria and Romania joined the European Union, and faced transitional provisions until 2014. In 2013, Croatia joined, and experienced labor market restrictions until 2020. Non-EU Member States (Iceland, Liechtenstein, Norway, and Switzerland[7]), also applied transitional provisions toward new EU Member States, hence we include them in our analysis.[8]

The option to unilaterally restrict labor markets generated different labor mobility patterns within the EU. For example, for 2004 entrants, Ireland, Sweden, and the UK decided to open their labor markets immediately in 2004 without any restric-

[6]EU15 denotes old EU Member States: Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, the Netherlands, Portugal, Spain, Sweden, and the United Kingdom.

[7]These four countries are denoted as +4.

[8]EU15+4 denotes all countries that applied transitional provisions.

tions, while other countries delayed or restricted access to specific industries. Denmark, Greece, Spain, and Portugal, for instance, removed restrictions only in 2009. France, Belgium, the Netherlands, and Austria gradually opened their labor markets, allowing only workers in certain industries and introducing quotas. Germany prohibited all access to the labor market, with the exception of a few sectors with labor shortages, until the expiry of the transitional provisions in 2011. Appendix Table A1 provides an overview of the precise opening dates for each destination (EU15+4) and origin (NMS) country.

Once a labor market has opened for immigrants, it is permanently open.[9] Therefore, this shock permanently changed migration patterns. The greatest increases in emigration rates, however, occurred in the years of the openings, when citizens in the NMS were allowed to work in the West without any legal restrictions for the first time.

One might argue that restricting a country's labor market is endogenous and related to local labor market conditions. Germany, for instance, experienced high unemployment rates during the mid-2000s and this was one of the reasons for its labor market restrictions. However, while the transitional provisions are endogenous to labor market conditions and firm productivity in the *destination* country, they are exogenous to firm outcomes in the countries of *origin*. Furthermore, restrictions needed to be the same for all NMS from a given accession year, which meant that origin-specific exceptions were not possible.

We thus use the information on labor mobility regulations within the EU to construct an instrument for the emigration of workers to circumvent the endogeneity of emigration. We obtain legal information from the Labor Reforms database (section on labor mobility) of the European Commission and complement it with information from national legislation.[10]

[9]There was one exception where Spain reintroduced restrictions for Romanian workers during the financial crisis.

[10]LABREF database, DG Employment, Inclusion and Social Affairs - European Commission

3.2 Migration Data and Descriptive Statistics

Our migration data is self-compiled industry-specific stock data from the national statistics offices of emigrants' *destination* countries (EU15+4).[11] For a few countries, it was not possible to obtain data from administrative sources, therefore we obtained data from the national labor force surveys or used other proxies (see Data Appendix A.3 for an overview of the precise data source by country).[12]

Figure 1 shows the stock of emigrants from NMS residing in EU15+4 countries (dark line). Emigration increased after the EU accession and subsequently remained at higher levels. For comparison, the light line shows the stock of emigrants from E15+4 residing in other EU15+4 countries, which remained stable throughout the time period of interest.

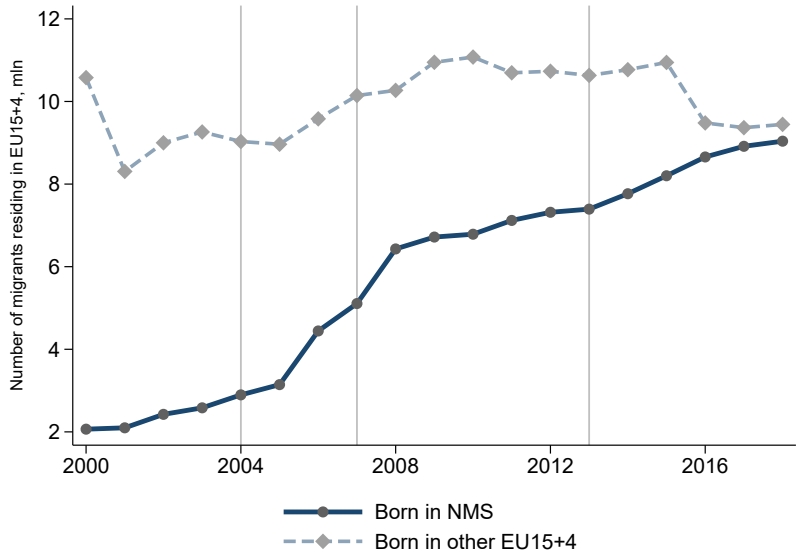
Kahanec, Zaiceva and Zimmermann (2009), Constant (2011), and Kahanec (2013) provide descriptive evidence of increased EU migration flows following the enlargements. Using country-level data, they show that the transitional provisions influenced migration movements. The UK and Ireland, for example, became the main EU destination country for Polish, Slovakian and Latvian workers. Kahanec, Pytliková and Zimmermann (2016) apply a difference-in-differences analysis and confirm that emigration from NMS increased with EU accession, but its full potential was hampered by the transitional provisions.

The self-selection of emigrants can be visualized by highlighting differences in skill levels between NMS emigrants and the NMS general population. Figure 2 uses the IAB Brain Drain Dataset developed by Brücker, Capuano and Marfouk (2013) and shows that on average highly skilled individuals are overrepresented in

[11] In general, administrative data on immigration are of better quality compared to data on emigration. First, not all emigrants officially report departure from their countries of origin. Second, even if they do, we are not likely to observe other characteristics, such as the industry of work prior to emigration.

[12] Other existing migration data sources could not be used as they do not include migration data at a two-digit industry level (this is the case for the global bilateral migration stock databases from the UN and the World Bank, the United Nations migration flows database), only provide observations every five or ten years or do not cover all the NMS (this is the case for the OECD DIOC data). The Eurostat Labor Force Survey aggregates all the relevant countries of origin into two groups (NMS10 for 2004-entrants and NMS3 for Bulgaria, Romania, and Croatia). Moreover, since it is a survey of around five percent of the population, some origin-destination-industry cells only have a few observations and are therefore unreliable.

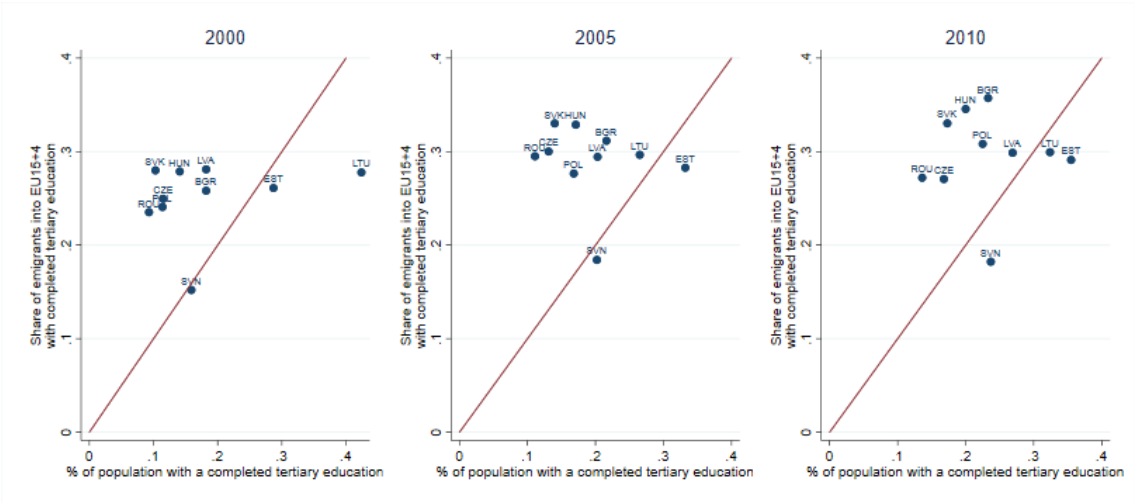
Figure 1: **Stock of EU Migrants Residing in EU15+4 Countries Before and After EU Accession**



Notes: This figure shows the evolution of emigration from NMS to EU15+4 and the evolution of migration within EU15+4. The y-axis indicates the number of migrants in millions. The dark (light) line shows the number of migrants residing in the EU15+4 from the NMS (EU15+4). Vertical lines correspond to the EU enlargements of 2004, 2007 and 2013.
Source: OECD Migration (MIG) Database.

the stock of emigrants from NMS living in EU15+4 compared to the share in the remaining population of their countries of origin. This indicates a loss of skilled human capital through emigration.

Figure 2: **Education Level of NMS Emigrants as Compared to the NMS General Population**



Notes: This figure shows the share of tertiary-educated emigrants in the emigrant stock of each of the NMS, contrasting it with the share of tertiary-educated individuals in the respective origin country. For most NMS, the share of tertiary educated is higher among emigrants than in the general population.
Sources: IAB Brain Drain Dataset (Brücker, Capuano and Marfouk, 2013), no data available after 2010. Eurostat for share of origin country population with tertiary education.

3.3 Firm Productivity Data and Descriptive Statistics

We obtain firm-level data from Bureau Van Dijk's Orbis and Amadeus databases, which provide standardized annual balance sheet and profit information for European public and private companies of all sizes.[13] We work with an unbalanced panel of approximately 230,400 manufacturing firms located in NMS. The period covered ranges from 2000 to 2016, with an average of about six annual observations for each firm. We only include companies from manufacturing in the sample because of the lower quality of firm data in other industries. In addition, productivity estimations in manufacturing are more concise than those in services, and firms are more comparable. Appendix A.3.3 provides a detailed description of the dataset and its coverage in NMS.

Our sample comprises firms with at least one employee and with available data to estimate TFP. Having examined the data, we also noted that a few firms report

[13] Orbis and Amadeus databases overlap to some extent. In our baseline specifications, we include firms available either in Orbis or Amadeus. Further in the text, we will use *Orbis* to denote this combined dataset.

unusually large figures for value-added and the number of employees, which is likely to be a reporting error. We, therefore, omitted these outliers in our baseline specifications (i.e., firms with value-added equal to or above the 99th percentile and firms with over 100,000 workers). Table 1 provides the summary statistics for firms in our sample.

Table 1: **Summary Statistics, Firm Data**

	Mean	SD	Min	Max	N
Firm age	10.250	8.502	0	319	1546053
Turnover, thousand EUR	1312.530	4334.518	0	1634908.1	1824455
Value added, thousand EUR	521.569	1221.195	0	10085.71	1825678
Total assets, thousand EUR	1222.553	16524.900	0	9217911.6	1800670
Fixed assets, thousand EUR	635.271	15207.310	0	8984836.5	1825678
Number of employees	30.177	141.309	1	56766	1825678
Material costs, thousand EUR	823.313	48451.803	0	65207768	1820111
Personnel costs, per employee, EUR	6448.540	20186.451	0	14183341	1825678
Total assets/L, EUR	66055.556	7.04e+05	0	3.372e+08	1800670
Fixed assets/L, EUR	32338.422	6.08e+05	0	3.355e+08	1825678
Y/L, EUR	24710.211	80179.337	0	9965801	1825678
Y/(WL)	8.272	196.622	0	124802.3	1824989
TFP (nace), LP	5.149	1.618	-12	15.018874	1725479
TFP (nace), Wooldridge	4.810	1.655	-12	15.022396	1725479

Notes: This table presents summary statistics for all firm-level variables from the Orbis sample used in our regression analysis. Y denotes value-added, L denotes number of employees, WL denotes total personnel costs. TFP is in natural logarithms. Firm age is calculated using the reported year of incorporation (nine firms in the sample turned out to be more than 200 years old, the oldest firms belong to the manufacturing of beverages and the manufacturing of machinery and equipment). The number of observations varies due to differences in the availability of variables.

We apply this firm-level data to calculate three measures of firm productivity: labor productivity, wage-adjusted labor productivity and TFP. The first measure is calculated as the total value-added over the number of employees. The second measure is calculated as the total value-added over personnel costs.[14] To obtain the third measure, we apply a semi-parametric approach as in Levinsohn and Petrin (2003) and Wooldridge (2009). This method enables us to overcome the simultaneity bias between firms' inputs and unobserved (to researchers) productivity shocks. To account for differences in the production function across sectors, we calculate TFP within two-digit NACE Revision 2 industries. For details regarding the TFP calculation, we refer to Appendix A.3.2. Other firm-level outcomes are the number of employees, personnel costs per employee, and the capital-labor ratio.

[14]Personnel costs include wages and other personnel-related expenditure, such as hiring and training costs.

In Appendix A.3.3 we provide details on the representativeness of Orbis data by comparing country-year aggregates obtained with data from Orbis and data from the Eurostat Structural Business Statistics (described below). As Figure A7 shows, Orbis coverage varies by country and over time. Since we use Orbis data to study the within-firm effects of emigration, our results are less sensitive to the changes in firm coverage by Orbis. We further check the robustness of our results by restricting our sample to those firms that existed prior to the EU enlargement. However, Orbis data is not well-suited to analyzing firms' entry and exit dynamics. Therefore, for the industry-level analysis we mainly resort to Eurostat data, which provides better coverage at the industry level.[15]

3.4 Additional Data

As additional covariates and for the analysis at the industry level, we use aggregated (two-digit NACE Revision 2) industry-level data that are available for all EU Member States and are harmonized by Eurostat. The Eurostat Structural Business Statistics database contains annual indicators, such as value-added, number of employees, and investment. We also use industry-level import data to control for competition faced by firms. These data come from the UN Comtrade database.[16]

We use additional data to illustrate the mechanisms. Industry-level data (two-digit NACE Revision 2) on training comes from the EU Labor Force Survey, an annual survey conducted in all EU Member States and compiled by Eurostat. We also utilize a measure of skill shortages from the EU Commission Business Survey, which is conducted in all EU member countries by the Directorate General for Economic and Financial Affairs (DG ECFIN). The survey addresses firms in the manufacturing, service, retail trade, and construction sectors and enquires about their assessment and expectations of business development. In addition to other questions, the survey's participants are asked to evaluate factors limiting their production (such as labor constraints or intense product-market competition). The EU

[15] To ensure that our firm-level results are not driven by data availability, we compare our industry-level results when using all available Eurostat data, and when restricting the sample to those countries and industries that are also well-covered in Orbis.

[16] The data are available at the product level, which we convert to the industry level using the NACE Revision 2 - NACE Revision 1 - HS 1996 correspondence table.

Commission publishes information on a two-digit NACE industry level. Thus, the measure obtained is equal to the share of firms in each industry reporting constraints due to labor or competition. Table A2 in the Appendix provides summary statistics for all additional variables used in the estimations.

4 Econometric Specification

Our baseline specification is a 2SLS regression of firm outcomes on the emigration rate using the "Free Labor Mobility" (FLM) variable as an instrument for emigration rates in countries of origin. This captures a local average treatment effect (LATE). The FLM variable, which we describe in Section 4.2 below, summarizes EU labor mobility regulations and quantifies the exposure of NMS firms to the emigration of their workforce. For comparison, we also run OLS regressions.

4.1 OLS Specification

We begin by estimating simple OLS regressions of firm outcomes on emigration. The regression equation we estimate is the following:

$$Y_{ft} = \alpha + \beta_1 Emigrants_{oit} + \beta_2 a_{ft} + \beta_3 a_{ft}^2 + \beta_4 I_{oit} + \tau_{ot} + \eta_{i't} + \nu_f + \varepsilon_{ft} \qquad (1)$$

Y_{ft} are outcomes of a firm (f) in year (t). The main dependent variables are three different measures of firm productivity. $Emigrants_{oit}$ denotes the emigration rate in a given firm's origin (o), industry (i) and year (t).[17] We use the contemporaneous value for the main independent variable because by construction we observe emigrants once they are already employed in the destination industry. β_1 captures the correlation between emigration and firm-level outcomes, controlling for a set of covariates and fixed effects. Variables a_{ft} and a_{ft}^2 account for firm age. I_{oit} includes origin-industry-specific linear time trends. Specifically, we calculate average growth rates in the origin-industry investment, employment, import share and

[17] It is defined as the stock of emigrants from a given origin-industry-year over the average number of employees in a given origin-industry as of 2000-2002 multiplied by 100.

value-added before 2004 and interact them with linear time trends.[18] τ_{ot} are origin-specific time dummies and $\eta_{i't}$ are industry-specific time dummies.[19]

ν_f represent firm fixed effects, and ε_{ft} is the error term. Standard errors are clustered at the origin-industry level, which allows us to account for common shocks not only at the level of the instrument's variation but also for common shocks across time.

In the baseline empirical model, we consider only within-firm variation. Such a specification allows us to abstract from the firm unobserved time-invariant heterogeneity (as initial management ability or quality of business ideas), and other constant characteristics of a firm's location or industry-specific production technologies. Origin-specific time fixed effects control for economic and policy shocks that concern all firms from the same origin country (e.g., introduction of the euro), while industry-specific time fixed effects account for possible changes that happen to all firms from the same industry (e.g., a technological or a price shock).

Notwithstanding, we encounter several problems with the OLS specification. First, we can face reverse causality if individuals leave sectors already experiencing negative productivity shocks. Second, there could be omitted variables, such as a change in an origin-industry-specific regulation, that can influence both emigration and firm productivity. Third, the main independent variable may suffer from measurement error. In the next section, we describe our instrumental variable approach.

4.2 Two Stage Least Squares Specification

Higher emigration of NMS workers was triggered by the opening of the EU15+4 labor markets. We capture these changes in EU labor mobility regulations by creating the Free Labor Mobility (FLM) variable, which we use as an instrument for

[18] We employ this specification to control for possible differential time trends, which could lead to differences in labor demand unrelated to emigration, and to avoid the 'bad controls' problem because the emigration of workers could have directly affected these industry characteristics.

[19] The industries i' are the following: food and beverages (Nace Revision 2 codes 10-12), textile (13-15), wood (16-18), coke and metals (19, 24-25), chemicals and pharmaceuticals (20-21), rubber and plastic (22-23), computers, electronics and machinery (26-28), transport items (29-30), other (31-33).

emigration in the 2SLS empirical specification. The first stage has the form:

$$Emigrants_{oit} = \alpha + \gamma_1 FLM_{oit} + \gamma_2 I_{oit} + \tau_{ot} + \eta_{i't} + \nu_{oi} + \phi_{oit} \quad (2)$$

Variables are denoted as above in Equation 1. We use Equation 2 at the origin-industry-year level (which is the level of variation in the FLM) to test alternative specifications of the instrument. We then estimate our main 2SLS regressions at the firm or industry level. Note that when performing the analysis at the firm level, the first stage is adjusted to the firm level.

We construct the instrument in three steps. First, for each origin-industry-year observation we obtain a set of 19 dummies D_{doit}, with each dummy corresponding to one of the EU15+4 destination countries, d. A dummy takes the value of one, if, according to an old EU Member State's regulation, a specific industry i is open to labor migrants from a given origin country o in year t. For example, the UK completely opened its labor market for the NMS (2004-entry) group in 2004. Therefore, UK dummies for every industry for all countries of origin from this accession year equal one starting from 2004. In contrast, France retained the transitional provisions for the 2004-entrants until 2008. Prior to 2008, the French government applied a special job scheme, authorizing free labor market access only in construction, tourism, and catering. In our sample, France dummies for NMS industries take a value of zero until 2008 for all manufacturing sectors.

Second, to account for different capacities of the destination labor markets to hire migrant workers, we complement the legislation dummies with the measures of the *pre-2004* potential labor market size. We use two alternative measures: Market size I - the number of employees in a destination industry multiplied by the share of firms reporting labor shortages; Market size II - the number of employees in a destination industry multiplied by the share of employed migrants from all origins. Reported labor shortages (or the share of employed migrants) proxy firms' willingness to hire foreign workers. To avoid endogeneity, all values (the number of employees, labor shortages, the share of employed migrants) to construct the potential market size for migrants are taken before 2004 - prior to the first labor

market opening.[20] We then interact Market size I (*or* Market size II) with the legal dummies to characterize the potential number of jobs for NMS migrants in a certain destination industry.

Third, to summarize the set of 19 potential labor markets into a single variable, we apply proximity weights reflecting how strongly the opening of a particular EU15+4 labor market affects citizens of a respective new Member State. Intuitively, the opening of the Finnish labor market should exert a stronger effect on emigrants from the Baltic states than on those from the Czech Republic or Slovenia. We apply and compare the performance of two alternative measures for such weights. One approach is based on the migration network literature. We use the previous (as of 2000) distribution of NMS migrants across EU15+4 countries: the larger the share of a given destination for emigration from a given origin, the larger the weight for a corresponding labor market.[21] This assumes that following the openings, labor migrants are more likely to go where previous migrants have gone. As an alternative proximity measure, we apply inverse bilateral distances between origin and destination countries.[22]

The instrument is described by the equation:

$$FLM_{oit} = \sum_{d=1}^{19} w_{do} \cdot D_{doit} \cdot MS_{di} \tag{3}$$

FLM_{oit} is the value for one observation (origin-industry-year). D_{doit} denotes the legal dummies, which can take the value of one if the labor market in a destination d industry i is open for the citizens of a given origin o in a given year t and the value zero otherwise. MS_{di} represents the potential labor market size in a destination d and industry i. w_{do} denotes the proximity weight. The FLM_{oit} variable can be interpreted as the potential number of jobs available to workers from origin o in year t and industry i across all EU15+4 countries. It has a value of zero for all origin industries before 2004 and then gradually increases as more and more destination

[20] For every variable, we take the average value over 2000-2003.

[21] To calculate the shares, we divide the stock of emigrants residing in a given destination as of 2000 by the origin's population as of 2000.

[22] We use population-weighted average distances between the most populated cities in origin and destination countries from the CEPII dataset Conte, Cotterlaz and Mayer (2022).

industries lift the labor mobility restrictions for NMS workers. For the regression analysis, we log-transform the FLM_{oit}.[23]

For the instrument to be valid, we need it to be relevant and to satisfy the exclusion restriction. The results for the first-stage regression are presented in the Appendix and corroborate the instrument's relevance: the larger the available pool of jobs in EU15+4, the larger the emigration rate is. Table A3 shows that our first-stage results are robust to using alternative measures of market size and proximity factors for constructing the FLM variable. For our baseline, we select the FLM specification that yields the largest F-Statistic. This specification uses previous migration networks to calculate proximity weights and the number of employees multiplied by the share of firms reporting labor shortages as the measure of the potential market size. Appendix Figures A1 and A2 illustrate the variation in the baseline FLM variable across time, industries and countries. Figure A3 plots emigration rates against the FLM variable (after partialling-out controls and fixed effects) to show that our variation is not driven by outliers.

Given that the FLM variable varies at the industry, origin, and year level, Table A3 illustrates how our instrument works with the industry-level data that covers all available industries in all NMS. When conducting the analysis at the firm level, the first stage is re-estimated with the available firm-level data, with the same fixed effects and controls as in the second stage. For completeness, we report the corresponding F-statistics, as well as the first-stage FLM coefficient and standard error under every 2SLS specification below.

An alternative to our instrument construction is to use a flexible non-parametric method (a random forest) to allow for various interactions and non-linearities in how legal dummies, market size and proximity weights shape emigration. This alternative specification of the instrument yields similar results. We present and discuss it in Section 7.

An important assumption for the relevance of the instrument is that emigrated workers stay in the same industry. The industry exemptions in the transitional provisions were broad (often defined at a two-digit industry level), and hence we consider it plausible that most skilled migrants stay in the same industry. If there were a lot

[23] We take a natural logarithm of $(FLM_{oit} + 1)$ to preserve zero values.

of industry switching, this would bias our estimates toward zero. While there is evidence that skilled immigrants are typically overqualified in the destination country (Drinkwater, Eade and Garapich, 2009; Visintin, Tijdens and van Klaveren, 2015; Chiswick and Miller, 2008, 2009; Johnston, Khattab and Manley, 2015; Lindley, 2009; Nielsen, 2011), they are still more likely to stay within the same sector. For instance, Kuhnen and Oyer (2016) show that firms try to reduce uncertainty regarding workers' future productivity and prefer to make offers to candidates who have experience in the firm's industry. To provide some evidence, we use EU LFS data where for a subset of NMS respondents we can observe both their industry of work in the country of origin and their current industry of work.[24] As Figure A4 in the Appendix shows, about 75 percent of NMS migrants stay in the same industry. This share is slightly higher among NMS migrants with lower or upper secondary degrees (close to 80 percent) and somewhat lower for NMS migrants with a tertiary degree who are also more likely to be overqualified (just below 70 percent).

Detailed industry-, year- and destination-specific labor mobility regulations were uncorrelated with other accession policy changes, such as the free movement of goods or capital, which entered into force either before or immediately upon the EU accession and concerned all industries. Since destination countries had to apply the same labor-market openings to all NMS from the same accession year, it is unlikely that a single firm (or even an industry from a certain origin) could have had an impact on these decisions or that an old EU Member State targeted a specific industry of a specific origin country. Table A4 checks for possible pre-trends and shows that future changes in the FLM variable (between 2004 and 2003, 2007 and 2003, and 2013 and 2003) do not predict prior industry-level changes (between 2003 and 2000) in emigration stocks, labor productivity, firm exits and the number of firms in NMS.

A possible threat to the exclusion restriction is that old EU Member States opened those industries earlier where they had received (or expected) a competitive advantage over their NMS rivals following the EU enlargement. In this case, however, one would already expect to see negative tendencies in NMS firm performance prior to the openings and emigration of workers. Moreover, we run a placebo

[24] Because of data restrictions, we can observe only 1-digit industry classification.

test to show that our instrument is not correlated with the perceived product-market competition in NMS (Table A19).

5 Results

This section presents and discusses the empirical results. Regressions in Subsection 5.1 and 5.2 include firm fixed effects and thus capture within-firm variation in outcomes as a response to the emigration of workers. Subsection 5.3 focuses on industry dynamics and thus takes the effects stemming from the exit and entry of firms into account.

5.1 Productivity Results

Table 2 presents our main results in Columns 4 - 6: we regress total factor productivity (Column 4), labor productivity (Column 5) and wage-adjusted labor productivity (Column 6) on the emigration rate, instrumented by the Free Labor Mobility (FLM) variable.[25] Columns 1-3 show respective OLS results for comparison. Firm productivity measures are in natural logarithms, the main independent variable (emigration rate) is in percentage points. Thus, coefficients can be interpreted as a percent change in firm productivity when the emigration rate changes by one percentage point.[26]

For an average firm in our sample, the causal effect of emigration on the three different measures of productivity is negative and significant. In the main specification, a one percentage point increase in emigration causes a decrease in firm productivity between 1.1 and 2.3 percent. For all three outcomes, the emigration coefficient is significant at the one-percent level.

There could be several reasons why our IV estimates are stronger than the OLS results. First, OLS estimates are likely to be biased toward zero due to measure-

[25] The obtained first-stage coefficients in the firm-level regressions are slightly different from those presented in Table A3 because of differences in the data coverage (some industries and countries of origin are not well represented in Orbis data) and differences in the empirical specification - using firm fixed effects and additional firm-level controls.

[26] The exact calculation for the effect of interest is $(e^{\beta_1} - 1) * 100$.

Table 2: **The Effect of Emigration on Firm Productivity**

	(1) OLS TFP	(2) OLS Y/L	(3) OLS Y/(WL)	(4) 2SLS TFP	(5) 2SLS Y/L	(6) 2SLS Y/(WL)
VARIABLES						
$Emigration_{oit}$	-0.001	-0.000	-0.000	-0.023	-0.014	-0.011
	(0.001)	(0.001)	(0.000)	(0.007)	(0.005)	(0.004)
Observations	1,473,412	1,473,412	1,473,412	1,473,412	1,473,412	1,473,412
R-squared	0.694	0.653	0.622			
Clusters	239	239	239	239	239	239
Mean outcome pre-2004	4.897	8.401	1.443	4.897	8.401	1.443
SD outcome pre-2004	1.496	1.457	0.786	1.496	1.457	0.786
Mean Δ emigration share	1.023	1.023	1.023	1.023	1.023	1.023
SD Δ emigration	6.860	6.860	6.860	6.860	6.860	6.860
Firm FE	✓	✓	✓	✓	✓	✓
Year#Origin FE	✓	✓	✓	✓	✓	✓
Year#Industry FE	✓	✓	✓	✓	✓	✓
Linear trends/controls	✓	✓	✓	✓	✓	✓
First stage F-stat				15.54	15.54	15.54
FLM coefficient				2.746	2.746	2.746
FLM se				0.697	0.697	0.697

Notes: The table presents OLS (Columns 1-3) and 2SLS (Columns 4-6) effects of emigration on three measures of firm productivity. Columns 1 and 4 show the effect of emigration on TFP (estimated according to the Levinson Petrin method). Columns 2 and 5 show results for labor productivity (value-added/employees) and Columns 3 and 6 show wage-adjusted labor productivity (value-added/employee costs). In addition to the fixed effects mentioned in the table, we control for differential time trends at the origin-industry level, by interacting average growth rate in employment, value-added, investment and import share between 2000 and 2003 with a linear time trend. Additional controls include firm age and age squared. Standard errors (in parentheses) are clustered at the origin-industry level.

ment error in the migration variable. Second, we proxy the stock of emigrants by the number of NMS-born workers employed in a respective industry of a EU15+4 country. This number may increase due to the emigration of workers, but also for other reasons: for instance, previously unemployed NMS migrants (from earlier emigration cohorts) or family migrants could become employed or former NMS students who had already studied in EU15+4 could enter the labor market there. While the emigration of workers is important for firms in countries of origin as they are losing workers, the other reasons should not fundamentally affect firms in countries of origin. When using the IV to predict emigration, by construction we are more likely to capture an increase due to labor emigration, since the identifying variation is generated by changes in labor mobility regulations.

One interesting question that arises is whether these productivity effects are temporary or permanent. To shed more light on this, we focus on firms that existed in 2003-2013 and estimate the emigration effect over this ten-year period. Table A5

illustrates that we still find an effect if we regress 10-year differences in the productivity measures (over 2013-2003) on 10-year differences in the emigration rate.[27] Results suggest that also in the longer term, productivity declines by between 1.6 percent for the wage-adjusted labor productivity and 4.6 percent for TFP when the emigration rate increases by one percentage point. For all three productivity outcomes, this effect is significant at the one-percent level. Both our contemporaneous results and our long-term results suggest that the decrease in TFP is stronger in magnitude than the decrease in both measures of labor productivity. This signals that firms' overall efficiency of inputs (not only of labor) decreases with emigration.

Our results are corroborated by looking at the reduced-form estimations (see Table A6 and Figure A5 in the Appendix). The reduced form abstracts from a measurement error that could arise in the migration data and shows how firm productivity changes with higher *exposure* to emigration. In addition, we can examine possible pre-trends. We provide results from two alternative specifications. Table A6 shows the results of a reduced-form regression where we include the contemporaneous value (FLM_{oit}), two lags (FLM_{oit-1}, FLM_{oit-2}), and two leads (FLM_{oit+1}, FLM_{oit+2}) of the FLM variable. We still find that an increase in our measure of free labor mobility significantly decreases firm productivity. Since the reduced form measures the effect of the law change and not the actual emigration, we see that firm productivity reacts with a time lag of one to two years. We also note that leads of the FLM variable do not have any significant impact on firm productivity.

Figure A5 refers to the results of an event-study specification. To capture variation over time, we center the data around the year of the first opening where period "zero" corresponds to the first year when FLM becomes positive for a given firm. To capture variation in the quantity of the treatment across firms, we interact the maximum FLM exposure that a firm faces in our sample with four time periods before and after the first opening. The sample, thus, includes firms that experience the first labor-market opening (FLM changes from zero to a positive value)

[27] Specifically, we take a difference between stocks of emigrants in 2013 and 2003 and divide it by the number of employees in a given origin-industry prior to 2004. In this way, we capture the cumulative net emigration rate between 2003 and 2013.

and that appear in the data for at least nine years. Other fixed effects and control variables remain the same as in the baseline specification. As Figure A5 shows, we do not observe strong pre-trends. At the same time, the effect of exposure to emigration following the first opening is negative and significant, and its magnitude increases over time. Note that in this specification we cannot disentangle what part of this dynamic effect is due to new openings and what part is due to the lagged firm responses. We discuss alternative event-study approaches based on the recent econometric literature and their limitations in our setting in Section 7.

5.2 Heterogeneity

In the main specification, we analyze the effect of higher emigration for the full sample of NMS firms. To examine heterogeneous effects, we repeat the estimations for different sub-samples of firms: 1) incumbents, 2) firms with different initial productivities.

We first look at incumbents, which are defined as firms that existed prior to the first EU enlargement in 2004. The advantage of focusing on these firms is that their productivity is neither affected by endogenous entry decisions nor by changes in data availability over time. Yet, as Table A7 shows, the results for incumbents do not differ significantly from the baseline results.

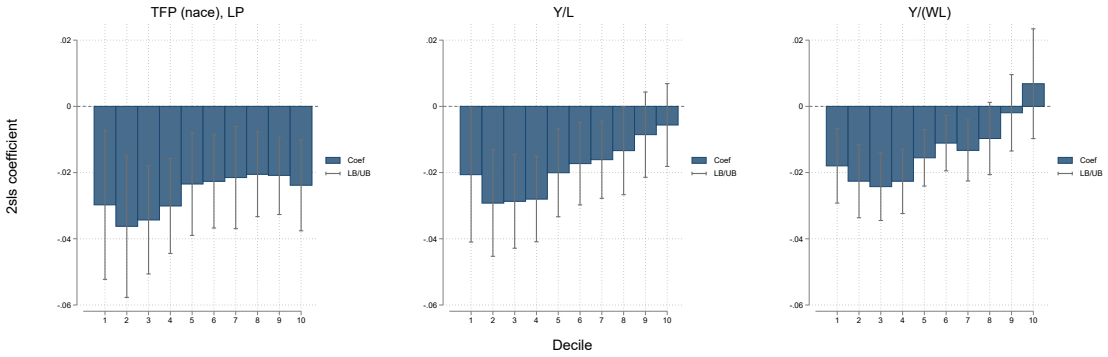
Once we estimate the effects for incumbent firms along different deciles of their initial productivity distribution, interesting patterns emerge. This analysis enables us to study whether initially more productive firms are differently affected by emigration compared to initially less productive firms.

Figure 3 highlights the contemporaneous effect of emigration on our three measures of firm productivity along the initial productivity distribution. To estimate the coefficients shown in this figure, the sample was split into deciles depending on a firm's total factor productivity prior to the first labor market opening.[28]

The effects of emigration on firm TPF are similar across all productivity deciles. However, for labor and wage-adjusted labor productivity, we observe that the negative emigration effect declines as we move along the productivity distribution from

[28] We focus on the incumbent firms and calculate their average TFP before 2004. We then group firms into deciles according to their initial TFP *within* their origin-industry.

Figure 3: **Results Along the Initial Productivity Distribution**



Notes: This figure presents regression results for incumbent firms along different productivity deciles. The sample only includes firms that existed prior to 2004 as firms are divided into productivity deciles based on their average productivity between 2000 and 2003. On the x-axes we indicate firm productivity deciles and on the y-axes we indicate the estimated coefficients from a 2SLS regression for firms from a given decile. The blue bars show the estimated coefficients for instrumented emigration rate and the black lines show the upper and lower bounds of the 95 percent confidence interval.

less to more productive firms. The emigration effects are not significantly different from zero for firms in the top two productivity deciles suggesting that these firms are able to better adapt to the emigration shock at least in terms of preserving the efficiency of their labor inputs. For less productive firms, the negative emigration effects are more pronounced: for instance, for firms in the four lowest productivity deciles, a one percentage point increase in the emigration rate translates into a 2-3 percent drop in their labor and wage-adjusted labor productivity and into a 3-3.5 percent drop in TFP.[29] We further discuss differences in adjustment to emigration across firms in Section 6.

5.3 Industry-level Results

We complement the firm-level analysis with results at the industry level. Thus, we can analyze the effect of emigration on aggregate productivity in the presence of firm entry and exit dynamics. Additionally, since we are using more complete Eurostat data, we can check whether the results differ when using all available coun-

[29]There is a somewhat puzzling effect on the productivity of firms in the lowest decile. We suggest that this is driven by stricter selection to survival for these firms: low-productivity firms are more likely to exit the market, and the fact that we still observe some of these firms in our sample several years later, means that they have managed to behave similarly to more productive firms.

tries of origin and industries, and when restricting the data to origins and industries covered by Orbis.

Table 3 shows that emigration increases firm exits (Column 1). A one percentage point increase in the emigration rate increases exits by about 7.7 percent. The result is significant at the ten-percent level. At the same time, we do not find any significant impact on the total number of firms (Column 2), suggesting that new firms enter and that emigration increases the dynamism and fosters creative destruction at the industry level. To further investigate this, we use Orbis data and denote firms that appear for the first time in our sample in a given year as "new". We then calculate the average TFP of these new firms in the first year of their existence to proxy for the productivity of entrants. As Column 6 shows, the total factor productivity of new firms in their first year of existence is indeed higher with the significance at the five-percent level in industries that are exposed to higher emigration rates.[30]

Columns 3-5 show that there is no negative emigration effect on productivity measures at the industry level. One reason for this is the exit of the presumably weakest firms and the entry of new stronger firms. Such exit and entry dynamic explains why the average TFP (Column 5) does not drop at the industry level.[31] Another reason is that the industry-level outcomes (Columns 3-4) give a larger weight to larger (and more productive) firms, which, according to the results presented in Section 5.2, experience a weaker effect of emigration on labor and wage-adjusted labor productivity. We can observe that the first-stage F-statistic is slightly higher at the industry level. This stems from better coverage of the industry-level data by Eurostat compared to the firm-level data by Orbis.[32] In addition, we exploit within-industry rather than (aggregated at the industry level) within-firm variation in the instrument.

We next look at the long-term changes in productivity at the industry level. Table A9 shows that emigration does not affect industry-level productivity in the

[30] We interpret the result in Column 6 with caution, however, due to changes in Orbis data coverage over time.

[31] We use again Orbis data and aggregate firm TFP at the origin-industry-year level. Note, the average TFP in Column 5 is not size-weighted.

[32] For example, in Orbis we have only a few observations from Hungary, Latvia and Lithuania that were strongly affected by emigration. However, the results are robust to restricting the sample to industries with good Orbis coverage as shown in Table A8.

Table 3: **Industry-level Effects, Eurostat Data (All Available Industries)**

	(1) 2SLS Exits	(2) 2SLS N firms	(3) 2SLS Y/L	(4) 2SLS Y/(WL)	(5) 2SLS Mean TFP	(6) 2SLS Mean TFP new firms
VARIABLES	Eurostat	Eurostat	Eurostat	Eurostat	Orbis	Orbis
$Emigration_{oit}$	0.074	-0.000	0.015	0.021	0.004	0.101
	(0.039)	(0.011)	(0.028)	(0.021)	(0.014)	(0.041)
Observations	4,105	4,021	3,782	3,765	3,819	3,819
Clusters	263	260	254	254	242	242
Mean outcome pre-2004	3.053	6.128	2.291	0.999	5.522	1.698
SD outcome pre-2004	2.135	1.934	0.653	0.256	1.071	1.787
Mean Δ emigration share	0.682	0.682	0.682	0.682	0.682	0.682
SD Δ emigration	5.541	5.541	5.541	5.541	5.541	5.541
First stage F-stat	17.16	21.70	13.49	13.77	13.06	13.06
FLM coefficient	1.253	1.439	1.253	1.288	1.161	1.161
FLM se	0.302	0.309	0.341	0.347	0.321	0.321
Origin#Industry FE	✓	✓	✓	✓	✓	✓
Year#Origin FE	✓	✓	✓	✓	✓	✓
Year#Industry FE	✓	✓	✓	✓	✓	✓
Linear trends/controls	✓	✓	✓	✓	✓	✓

Notes: The table presents 2SLS regressions of different industry-level outcomes on emigration. Column 1 shows the number of firm exits and Column 2 the number of firms. Columns 3-6 show measures for productivity. Columns 3 and 4 use Eurostat data: Y/L is calculated by dividing the total value-added on the total number of workers, $Y/(WL)$ is calculated by dividing the total value-added on the total employee costs. Column 5 uses Orbis data to calculate the average TFP of all present firms in a given origin-industry-year. Column 6 uses Orbis data to calculate the average first-year TFP of entering firms. In addition to the fixed effects mentioned in the table, we control for differential time trends at the origin-industry level by interacting average growth rate in employment, value-added, investment and import share between 2000 and 2003 with a linear time trend. Standard errors (in parentheses) are clustered at the origin-industry level. Note that the number of observations and hence the first-stage statistics in Columns 1-6 vary due to differences in data availability. The results hold when we restrict the sample to industries with no missing data for all six outcomes.

long run. If anything, labor, wage-adjusted labor productivity and the average TFP slightly increase, but these effects are not statistically significant. There is suggestive evidence (Column 5) for labor reallocation toward more productive firms: the average size-weighted TFP is higher in industries with higher exposure to emigration. This effect is significant at the ten-percent level. Since the average size-weighted TFP for always-present firms between 2003 and 2013 is not affected by emigration (Column 6), we can conclude that the main reallocation takes place from the exiting firms toward more productive stayers and new firms.

6 The Mechanisms

To shed light on how firms adjust to emigration, we look at three other firm-level outcomes: personnel costs, the capital-labor ratio and the number of employees, as well as at two additional industry-level outcomes: the percentage of firms indicating skill shortages and the percentage of employees receiving training.

Column 1 in Table 4 shows that emigration does not change personnel costs. The measure we use includes wages, as well as other employee-related expenses, such as hiring and training costs. The regression result captures different margins of adjustment within a firm, which seem to cancel each other out. On the one hand, wages should increase because labor becomes scarcer. On the other hand, firms can replace more experienced and therefore also more expensive emigrated workers with new (less experienced and less skilled) hires who have lower wages.[33] In line with this substitution effect, we find that the number of employees does not fall in response to emigration. In our baseline specification (Column 3), we obtain a small positive but insignificant effect on the number of employees.

Column 2 in Table 4 provides evidence that firms downward adjust their capital-labor ratio in response to emigration. This lack of investment is persistent in the longer term (see Table A10 for an estimation over a ten-year period). A possible explanation is that firms treat capital and skilled labor as complements. Firms may reduce their investments if they anticipate difficulties with retaining and hiring qualified workers. Similarly to the productivity results (Table 1), we observe that the emigration effect on the capital-labor ratio is driven by less productive firms, while it is not significantly different from zero for firms in the top two productivity deciles (Figure A6). It could be that high-productivity firms are better able to substitute those who left with workers of similar quality, hence we observe no effects on both their labor productivity and capital-labor ratio. As replacement workers have less firm-specific skills and need time to be trained, we still find a negative

[33]The unemployment rate in most new EU Member States had constituted more than ten percent just before their EU accession and then decreased substantially. In Poland, for instance, the unemployment rate reached 20 percent in 2003 and dropped to eleven percent by the end of 2006 (Eurostat *une_rt_m* dataset). Among other factors, better emigration opportunities could have contributed to the decrease in unemployment both directly - by letting individuals find jobs abroad, and indirectly - by releasing jobs previously held by employed emigrants.

effect on firm TFP even for the most productive firms.

Table 4: **The Effect of Emigration on Other Firm Outcomes**

VARIABLES	(1) 2SLS Costs per employee	(2) 2SLS Assets/L	(3) 2SLS N employees
$Emigration_{oit}$	-0.000	-0.020	0.008
	(0.002)	(0.005)	(0.005)
Observations	1,473,412	1,473,412	1,473,412
Clusters	239	239	239
Mean outcome pre-2004	7.360	8.945	2.533
SD outcome pre-2004	1.168	1.517	1.424
Mean Δ emigration share	1.023	1.023	1.023
SD Δ emigration	6.860	6.860	6.860
First stage F-stat	15.54	15.54	15.54
FLM coefficient	2.746	2.746	2.746
FLM se	0.697	0.697	0.697
Firm FE	✓	✓	✓
Year#Origin FE	✓	✓	✓
Year#Industry FE	✓	✓	✓
Linear trends/controls	✓	✓	✓

Notes: The table presents 2SLS effects of emigration on various firm outcomes. Column 1 shows results for costs per employee. This includes wages and other labor costs. Column 2 shows results for assets over the number of employees, which we use to measure the capital-labor ratio. Column 3 shows results for the number of employees. In addition to the fixed effects mentioned in the table, we control for differential time trends at the origin-industry level, by interacting average growth rate in employment, value-added, investment and import share between 2000 and 2003 with a linear time trend. Additional controls include firm age and age squared. Standard errors (in parentheses) are clustered at the origin-industry level.

To further illustrate a link between emigration and human capital constraints faced by firms, we apply a measure of skill shortages available in the regular EU Commission Business Survey. The survey collects harmonized data from representative samples of firms in each EU Member State. In particular, responding firms are asked whether labor constraints represent a factor limiting their production. The survey providers then aggregate the answers at the origin-industry-year level. Column 1 in Table 5 shows that firms in industries that experience higher emigration instrumented by the FLM variable are more likely to report skill shortages. A one percentage point increase in the emigration rate leads to a 0.92 percentage point increase in the share of firms reporting skill shortages (significant at the five-percent confidence level). From Table A2, this corresponds to a seven percent increase in the reported skill shortages. Economists are often skeptical when listening to firms' concerns about skill shortages, as they can be alleviated by paying higher wages

to attract workers.[34] In our setting, however, the reported skill shortages also reflect the fact that even when firms manage to replace emigrated workers with new workers, those new workers are less productive. This could be due to the lack of firm-specific knowledge and skills, which are not accounted for in wages.

Firm-specific human capital can be accumulated either through learning by doing or through firm-specific training. Columns 2 and 3 in Table 5 provide evidence that firms invest less in general training and job-related training once the emigration opportunities of their (potential) workers and hence the expected turnover increase. The coefficients are negative, however, imprecisely estimated.

We acknowledge that due to the limitations of our data, we cannot directly test the mechanisms outlined above and treat the provided evidence as suggestive. We do not have perfect data to disentangle the exact channels in play. Therefore, better data (such as skill composition at the firm level, measures for firm-specific human capital, detailed migration data by industry and skill level, and detailed training and tenure data at the firm level) and further research are necessary to analyze the mechanisms with more precision.

[34] If it was only the wage increase that were the case, we would not have observed the drop in firm TFP.

Table 5: **Effect on Skill Shortages and Training**

VARIABLES	(1) 2SLS Skill shortages	(2) 2SLS Had training	(3) 2SLS Job-related training
Migrants	0.920	-0.165	-0.237
	(0.427)	(0.199)	(0.215)
Observations	3,281	3,144	3,144
Clusters	244	248	248
Mean outcome pre-2004	11.03	0.795	0.738
SD outcome pre-2004	11.57	2.473	2.393
Mean Δ emigration share	0.682	0.682	0.682
SD Δ emigration	5.541	5.541	5.541
First stage F-stat	13.84	7.318	7.318
FLM coefficient	1.059	1.164	1.164
FLM se	0.285	0.430	0.430
Origin#Industry FE	✓	✓	✓
Year#Origin FE	✓	✓	✓
Year#Industry FE	✓	✓	✓
Linear trends/controls	✓	✓	✓

Notes: Columns 1, 2 and 3 present industry-level 2SLS regressions of skills shortages, general training and job-related training on instrumented emigration. The variable skill shortages measures the percentage of firms in a given industry, country and year that indicates being constrained in production by a lack of skilled workers. The general training variable measures the percentage of employees that received training. The variable job-related training measures the percentage of employees that received job-related training. In addition to the fixed effects mentioned in the table, we control for differential time trends at the origin-industry level, by interacting average growth rate in employment, value-added, investment and import share between 2000 and 2003 with a linear time trend. Standard errors (in parentheses) are clustered at the origin-industry level.

7 Robustness

7.1 Changes in the Sample Composition

The results are robust to sample modifications. First, we exclude very small and very large firms to show that results are not driven by outliers in terms of firm size. Appendix Table A11 shows the main results when the sample of firms is restricted to small and medium-sized firms (with 5 - 250 employees). While the sample size drops from almost 1.5 million firm-year observations to less than a million, all our main results still hold. This enables us to conclude that the results are not driven by very small or very large firms behaving differently.

The results are also robust to excluding the years of the financial crisis 2009-2011. One concern could be that NMS firms were hit hard by the crisis, and workers had to move to EU15+4 in search of better opportunities abroad. If this coincided with specific labor-market openings, it could bias our estimates. However, for the

sample that excludes the observations between 2009 and 2011, all productivity indicators are still negative and significant at least at the five-percent level with the size of coefficients similar to the baseline specification as shown in Appendix Table A12.

We check whether the results are driven by certain countries. Table A13 shows the results when excluding the country with the largest number of observations (Romania), as well as excluding Hungary, Lithuania and Latvia, which are not well covered in Orbis and might constitute a selected sample of firms. Even though the number of observations drops strongly from 1.5 million to about 800,000 after excluding these four countries, we still find negative and significant effects on firm productivity.

7.2 Changes in the Empirical Specification

There are different ways to calculate TFP. We follow the Levinsohn-Petrin method (Levinsohn and Petrin, 2003) in our baseline specification. Another widely used calculation method is the Wooldridge method (Wooldridge, 2009). Results are very similar as can be seen in Table A15. Details on TFP calculations can be found in the Appendix A.3.2.

We run two additional robustness checks to check whether the results are robust to changes in the empirical specification. First, we add firm-specific linear time trends in addition to firm fixed effects.[35] Table A16 shows that results are robust. We then analyze whether the results are sensitive to the inclusion of control variables and run the regression without any controls. Table A17 illustrates that results are very similar if we keep only fixed effects.

7.3 Robustness of the Instrument

The FLM variable measures the exposure to emigration at the origin-industry-year level and is constructed by aggregating destination-origin-industry-year legal dummies D_{doit} multiplied by the potential market size for migrants in destination MS_{di}

[35] We focus on the sample of incumbent firms only and interact their average growth in TFP and employment before 2004 with a linear time trend.

and bilateral proximity weights w_{do} (Equation 3). Thus, we assume that the bilateral (origin-destination) industry-level emigration is proportional to the interaction: $w_{do} \cdot D_{doit} \cdot MS_{di}$.

We employ an alternative method to construct the instrument. Instead of using Equation 3 to construct the FLM variable and use it as an instrument for emigration at the *origin*-industry-year level, we first model emigration stocks at the *destination-origin*-industry-year level using components of the FLM variable and then aggregate the obtained predictions at the origin-industry-year level to construct an alternative instrument. The advantage of this method is that we can use flexible non-parametric methods (e.g. a random forest) to allow for various interactions and non-linearities in how the independent variables (legal dummies, market size, and proximity) shape emigration. Moreover, in this specification, we can combine alternative measures for the market size and proximity weights in one estimator.

To construct an alternative instrument, we start with a bilateral migration dataset at the destination-origin-industry-year level. These data characterize the stock of emigrants from origin o working in destination d, industry i in year t. First, we partial-out destination-origin-industry, destination-year, origin-year and industry-year fixed effects.[36] Second, we use components of the FLM variable to model residualized bilateral migration. We include legal dummies, two proximity measures (migration networks and distance), labor-market size, the share of firms with skill shortages and the share of employed migrants to the total employment in the destination industries. As in the baseline IV, only legal dummies vary over time, while other independent variables are fixed prior to 2004.[37]

We then apply a non-parametric ensemble method - random forest - to model the data.[38] The model fit ($R-squared$) constituted 0.21 in-sample and 0.032 out-

[36]We focus on residuals for two reasons. First, we employ these fixed effects in our baseline regression models. Second, non-parametric models do not perform well when the number of parameters is high relative to the number of observations. Moreover, non-parametric models assume independence between observations, which is not the case in panel data. Note that we can only partially address this issue by removing the above fixed effects.

[37]We repeated the exercise by including only those parameters that we use in baseline IV: legal dummies, migration networks, labor market size and skill shortages. The fit was slightly worse, but qualitatively the results remained similar.

[38]The random forest approach constructs a multitude of decision trees (400 in our case). Each decision tree takes a random sample with replacement from the dataset and conducts sequential

of-sample for the full model.[39] We can also examine the relative importance of our parameters for the model fit.[40] In our case, proximity weights (migration networks and distance) show similar performance: each of them accounts for about 21-22 percent of the total variation. They are followed by the market size parameters (number of employees in an industry - 18 percent, the share of employed migrants - 16.5 percent and skill shortages - 14 percent). Legal dummies generate about eight percent of the variation in the residualized bilateral migration. We obtain the alternative instrument for the emigration rate by further aggregating the out-of-bag predictions at the origin, industry, and year level.

Table A18 presents the results of the 2SLS regression with the random-forest instrument. While we do not observe an improvement over the baseline instrument in terms of the first stage, all main results hold.

7.4 Placebo Test

The placebo test uses an outcome that should not be affected by changes in the FLM, but that also varies at the origin-industry-year level and could capture economic or regulatory changes: firms' perception of the product-market competition. In the EU Commission Business Survey, besides skill shortages, firms also report on business challenges arising from stronger competition. Appendix Table A19 presents first-stage regression results with competition as the dependent variable. While all IV modifications returned statistically significant coefficients with high F-statistics for the emigration rate, none of them is correlated with the perceived

splitting of the data into groups (nodes) based on the values of the parameters (our independent variables). The model prediction from a single decision tree is the average of observations that appear in the same final node. The algorithm splits the data until further splitting no longer adds value to the prediction or a stopping rule is achieved (in our case, there should be at least five observations in each final node). The random forest then averages predictions of the individual decision trees. One advantage of the ensemble method is that only a subset of observations is used in the estimation of each tree within a forest (due to the sampling with replacement). Hence, one can construct an *out-of-bag* prediction for each observation by averaging predictions from only those trees that do not contain this observation when estimating the model. This should reduce the risk of overfitting.

[39] Out-of-sample fit in a random-forest model was proxied with an out-of-bag score.

[40] It is calculated for each parameter as the share of data splits across this parameter to the total number of splits. The idea is that there will be more splits across those parameters that are responsible for more variation in the dependent outcome.

competition. This result confirms that the constructed IV captures the exposure to higher emigration as opposed to other contemporaneous shocks.

7.5 New econometric difference-in-differences methods

Recently, new econometric methods to estimate heterogeneity-robust difference-in-differences estimators allowing dynamic effects have been proposed by Callaway and Sait'Anna (2021), Chaisemartin and D'Haultfoeuille (2021a), Borusyak et al. (2021), Sun and Abraham (2021) and Callaway et al. (2021). We have considered using these advanced econometric methods but unfortunately they cannot be applied in our context. We cannot apply Callaway and Sait'Anna (2021), Borusyak et al. (2021) and Sun and Abraham (2021) because their method only works for a binary treatment and we have a continuous treatment. One advantage of our setting is that by using the classic 2SLS, we can exploit a more detailed continuous treatment, which is not possible in the method mentioned above. The method by de Chaisemartin and D'Haultfoeuille (2021a) and Callaway et al. (2021) allow for continuous and staggered treatment. However, these methods require a number of very strong assumptions on alternative parallel trends (for all doses of treatment), which we cannot ensure. In addition, in our case, there are multiple time periods, variation in treatment timing and in treatment intensity, which make these regressions additionally sensitive to (i) treatment effect dynamics and (ii) heterogeneous causal responses across timing groups. We therefore abstain from using these methods that come with several drawbacks in our setting and focus on a more straightforward IV design.

8 Conclusion and Policy Implications

Countries in Central and Eastern Europe have experienced a large emigration wave following the accession to the EU. While emigration has economically benefited the individual migrants and the destination countries in most cases, the effects on the countries of origin are more contentious. This paper uses firm- and industry-level

panel data to illustrate a negative causal effect of emigration on the productivity of an average firm. To overcome the endogeneity problem, we exploit the natural experiment of EU enlargements and show how the gradual and industry-specific opening of the destination labor markets has created exogenous variation in emigration experienced by countries of origin. We show that the emigration of workers resulted in lower productivity for individual firms in Central and Eastern Europe and that the effect is persistent. This effect can be observed for most firms along the initial productivity distribution, except for the most productive ones. At the industry level, however, emigration leads to the exit of the weakest firms and the entry of new and more productive firms, thus promoting creative destruction and, as the result, mitigating the negative impact of emigration.

Our results are important both for firms and policymakers. Being aware of the challenge of emigrating workers helps firms react in a timely and adequate way. Firms may benefit from investing in automation technology or active human resource strategies, focusing, for instance, on providing training and retention measures. For policymakers, the effects of emigration "are not a matter of fate, [but] to a large extent, they depend on the public policies adopted in the destination and countries of origin".[41] We see three areas of policy interventions.

First, policymakers could enhance their efforts to increase the labor force participation of the existing population through investments in education and the encouragement of women. The prevalence of skill shortages justifies the need to invest in the skills of the local labor force and to mitigate search frictions. A skill upgrading of the local labor force could be addressed in the short-term by providing specific training courses, which the state could subsidize to alleviate the burden on firms. Over the long term, the education system should be better adjusted to labor market needs. Knowing that those skilled people are needed, may justify the investment. An increase in local human capital might also take place over the long term due to increased incentives to invest in education, which increase with the prospect to emigrate (Beine, Docquier and Rapoport, 2001). While a small fraction of the population will in fact emigrate, a significant fraction of well-educated workers typically remains and can help to develop the country. Further leeway would be to encourage

[41] Docquier and Rapoport (2012).

more women to participate in the labor force by improving access and quality of childcare, and abolishing disincentives to work for the second earner (Atoyan et al., 2016).

Second, policymakers could encourage return migration and immigration from other countries to increase the skilled workforce. Return migration may have various benefits for the origin country. Bahar et al. (2022) show that former Yugoslavia benefited from return migration, which boosted their exports. A few countries such as Poland and Lithuania have already recognized this potential and introduced incentives to encourage return migration. Even if return migration is low, firms could benefit from knowledge transfers, if firms and policymakers succeed in maintaining close ties with the diaspora (Fackler, Giesing and Laurentsyeva, 2020). Another opportunity is to attract and integrate immigrants and refugees from other Eastern European countries such as Belarus or Ukraine. Between 2015 and 2021, Poland, for instance, has provided more than half a million work and residence permits each year for workers from Ukraine and Belarus and since 2022 it has hosted more than 1.5 million Ukrainian refugees. By supporting their labor market integration (Battisti, Giesing and Laurentsyeva, 2019), policymakers could alleviate the skill shortages of local firms.

Third, the fact that certain European countries benefit, and others could lose from free labor mobility, justifies channelling EU structural and cohesion funds to countries that bore the brunt of emigrants' education expenses, while not benefiting from them.

References

Anelli, Massimo, Gaetano Basso, Giuseppe Ippedico, and Giovanni Peri. 2022. "Emigration and entrepreneurial drain." *American Economic Journal: Applied Economics. Forthcoming.*

Atoyan, Ruben V, Lone Engbo Christiansen, Allan Dizioli, Christian H Ebeke, Nadeem Ilahi, Anna Ilyina, Gil Mehrez, Haonan Qu, Faezeh Raei, Alaina P Rhee, and Daria V Zakharova. 2016. "Emigration and Its Economic Impact on Eastern Europe." *Staff Discussion Notes*, 2016(007): A001.

Bahar, Dany, Andreas Hauptmann, Cem Özgüzel, and Hillel Rapoport. 2022. "Migration and Knowledge Diffusion: The Effect of Returning Refugees on Export Performance in the Former Yugoslavia." *The Review of Economics and Statistics*, 1–50.

Bajgar, Matej, Giuseppe Berlingieri, Sara Calligaris, Chiara Criscuolo, and Jonathan Timmis. 2020. "Coverage and representativeness of Orbis data." OECD Science, Technology and Industry Working Papers, No.2020/06.

Bartelsman, Eric, John Haltiwanger, and Stefano Scarpetta. 2013. "Cross-Country Differences in Productivity: The Role of Allocation and Selection." *American Economic Review*, 103(1): 305–34.

Battisti, Michele, Yvonne Giesing, and Nadzeya Laurentsyeva. 2019. "Can job search assistance improve the labour market integration of refugees? Evidence from a field experiment." *Labour Economics*, 61: 101745.

Beerli, Andreas, Jan Ruffner, Michael Siegenthaler, and Giovanni Peri. 2021. "The abolition of immigration restrictions and the performance of firms and workers: Evidence from Switzerland." *American Economic Review*, 111(3): 976–1012.

Beer, Sebastian, and Jan Loeprick. 2015. "Profit shifting: drivers of transfer (mis)pricing and the potential of countermeasures." *International Tax and Public Finance*, 22(3): 426–451.

Beine, Michel, Frédéric Docquier, and Hillel Rapoport. 2001. "Brain drain and economic growth: theory and evidence." *Journal of Development Economics*, 64(1): 275–289.

Beine, Michel, Pauline Bourgeon, and Jean-Charles Bricongne. 2019. "Aggregate Fluctuations and International Migration." *The Scandinavian Journal of Economics*, 121(1): 117–152.

Blalock, Garrick, and Paul J Gertler. 2004. "Learning from exporting revisited in a less developed setting." *Journal of Development Economics*, 75(2): 397–416.

Bloom, Nicholas, and John Van Reenen. 2007. "Measuring and Explaining Management Practices Across Firms and Countries." *The Quarterly Journal of Economics*, 122(4): 1351–1408.

Bloom, Nicholas, Raffaella Sadun, and John Van Reenen. 2012. "The Organization of Firms Across Countries." *The Quarterly Journal of Economics*, 127(4): 1663–1705.

Bollard, Albert, David McKenzie, Melanie Morten, and Hillel Rapoport. 2011. "Remittances and the Brain Drain Revisited." *World Bank Economic Review*, 25(1): 132–156.

Brücker, Herbert, Stella Capuano, and Abdeslam Marfouk. 2013. "Education, gender and international migration: insights from a panel-dataset 1980-2010." *mimeo*.

Caliendo, Lorenzo, Luca David Opromolla, Fernando Parro, and Alessandro Sforza. 2021. "Goods and factor market integration: a quantitative assessment of the EU enlargement." *Journal of Political Economy*, 129(12): 3491–3545.

Chiswick, Barry R., and Paul W. Miller. 2008. "Why Is the Payoff to Schooling Smaller for Immigrants?" *Labour Economics*, 15(6): 1317 – 1340.

Chiswick, Barry R., and Paul W. Miller. 2009. "The International Transferability of Immigrants' Human Capital." *Economics of Education Review*, 28(2): 162 – 169.

Clemens, Michael A. 2013. "Why Do Programmers Earn More in Houston Than Hyderabad? Evidence from Randomized Processing of US Visas." *American Economic Review*, 103(3): 198–202.

Constant, Amelie F. 2011. "Sizing It Up: Labor Migration Lessons of the EU Enlargement to 27." Institute for the Study of Labor (IZA) IZA Discussion Papers 6119.

Conte, Maddalena, Pierre Cotterlaz, and Thierry Mayer. 2022. "The CEPII Gravity Database Highlights." *CEPII: Centre d'Etudes Prospectives et d'Informations Internationales*.

Dicarlo, Emanuele. 2022. "How do Firms Adjust to Negative Labor Supply Shocks? Evidence from Migration Outflows." Institute for the Study of Labor (IZA) IZA Discussion Papers 14994.

Docquier, Frédéric, and Hillel Rapoport. 2008. "Skilled migration: the perspective of developing countries." Institute for the Study of Labor (IZA) IZA Discussion Papers 2873.

Docquier, Frédéric, and Hillel Rapoport. 2012. "Globalization, Brain Drain, and Development." *Journal of Economic Literature*, 50(3): 681–730.

Docquier, Frédéric, Olivier Lohest, and Abdeslam Marfouk. 2007. "Brain Drain in Developing Countries." *World Bank Economic Review*, 21(2): 193–218.

Docquier, Frédéric, Çağlar Ozden, and Giovanni Peri. 2014. "The Labour Market Effects of Immigration and Emigration in OECD Countries." *The Economic Journal*, 124(579): 1106–1145.

Drinkwater, Stephen, John Eade, and Michal Garapich. 2009. "Poles Apart?

EU Enlargement and the Labour Market Outcomes of Immigrants in the United Kingdom." *International Migration*, 47(1): 161 – 190.

Dustmann, Christian, and Albrecht Glitz. 2015. "How Do Industries and Firms Respond to Changes in Local Labor Supply?" *Journal of Labor Economics*, Vol. 33 (3).

Dustmann, Christian, Tommaso Frattini, and Anna Rosso. 2015. "The Effect of Emigration from Poland on Polish Wages." *Scandinavian Journal of Economics*, 117(2): 522–564.

Dustmann, Christian, Uta Schönberg, and Jan Stuhler. 2017. "Labor Supply Shocks, Native Wages, and the Adjustment of Local Employment." *The Quarterly Journal of Economics*, 132(1): 435–483.

Egger, Peter, Simon Loretz, Michael Pfaffermayr, and Hannes Winner. 2009. "Firm-specific forward-looking effective tax rates." *International Tax and Public Finance*, 16(6): 850–870.

Elsner, Benjamin. 2013. "Does emigration benefit the stayers? Evidence from EU enlargement." *Journal of Population Economics*, 26(2): 531–553.

Fackler, Thomas A., Yvonne Giesing, and Nadzeya Laurentsyeva. 2020. "Knowledge remittances: Does emigration foster innovation?" *Research Policy*, 49(9): 103863.

Fox, Jeremy T., and Valerie Smeets. 2011. "Does Input Quality Drive Measured Differences In Firm Productivity?" *International Economic Review*, 52(4): 961–989.

Freeman, Richard B. 2006. "People Flows in Globalization." *Journal of Economic Perspectives*, 20(2): 145–170.

Gal, Peter N. 2013. "Measuring Total Factor Productivity at the Firm Level using OECD-ORBIS." OECD Economics Department Working Papers, No.1049.

Gorodnichenko, Yuriy, and Monika Schnitzer. 2013. "Financial Constraints And Innovation: Why Poor Countries Don't Catch Up." *Journal of the European Economic Association*, 11(5): 1115–1152.

Grogger, Jeffrey, and Gordon H. Hanson. 2011. "Income maximization and the selection and sorting of international migrants." *Journal of Development Economics*, 95(1): 42–57.

Grossmann, Volker, and David Stadelmann. 2011. "Does international mobility of high-skilled workers aggravate between-country inequality?" *Journal of Development Economics*, 95(1): 88–94.

Grossmann, Volker, and David Stadelmann. 2013. "Wage Effects of High-Skilled Migration: International Evidence." *World Bank Economic Review*, 27(2): 297–319.

Hafner, Flavio. 2021. "Labor market competition, wages and worker mobility." Pompeu Farba University Working Paper.

Hunt, Jennifer. 2011. "Which Immigrants Are Most Innovative and Entrepreneurial? Distinctions by Entry Visa." *Journal of Labor Economics*, 29(3): 417 – 457.

Hunt, Jennifer. 2015. "Are Immigrants the Most Skilled US Computer and Engineering Workers?" *Journal of Labor Economics*, 33(S1): S39 – S77.

Imbert, Clement, Marlon Seror, Yifan Zhang, and Yanos Zylberberg. 2022. "Migrants and firms: Evidence from china." *American Economic Review*, 112(6): 1885–1914.

Jäger, Simon. 2016. "How Substitutable Are Workers? Evidence from Worker Deaths." Working Paper.

Johnston, Ron, Nabil Khattab, and David Manley. 2015. "East versus West? Over-qualification and Earnings among the UK's European Migrants." *Journal of Ethnic and Migration Studies*, 41(2): 196–218.

Kahanec, Martin. 2013. "Labor mobility in an enlarged European Union." *International Handbook on the Economics of Migration*. Cheltenham, UK:Edward Elgar Publishing.

Kahanec, Martin, Anzelika Zaiceva, and Klaus F Zimmermann. 2009. "Lessons from migration after EU enlargement." In *EU labor markets after post-enlargement migration*. 3–45. Springer.

Kahanec, Martin, Mariola Pytliková, and Klaus F. Zimmermann. 2016. "The Free Movement of Workers in an Enlarged European Union: Institutional Underpinnings of Economic Adjustment." *Labor Migration, EU Enlargement, and the Great Recession*, , ed. Martin Kahanec and Klaus F. Zimmermann, 1–34. Berlin, Heidelberg:Springer Berlin Heidelberg.

Kapur, Devesh, and John McHale. 2005. *Give Us Your Best and Brightest: The Global Hunt for Talent and its Impact on the Developing World.* Center for Global Development.

Kerr, Sari Pekkala, and William R. Kerr. 2013. "Immigration and Employer Transitions for STEM Workers." *American Economic Review*, 103(3): 193–97.

Kerr, Sari Pekkala, William R. Kerr, and William F. Lincoln. 2014. "Firms and the Economics of Skilled Immigration." In *Innovation Policy and the Economy, Volume 15*. *NBER Chapters*. National Bureau of Economic Research, Inc.

Kerr, William R. 2013. "U.S. High-Skilled Immigration, Innovation, and Entrepreneurship: Empirical Approaches and Evidence." National Bureau of Economic Research, Inc NBER Working Papers 19377.

Kuhnen, Camelia M, and Paul Oyer. 2016. "Exploration for human capital: Evidence from the MBA labor market." *Journal of Labor Economics*, 34(S2): S255–S286.

Levinsohn, James, and Amil Petrin. 2003. "Estimating Production Functions Using Inputs to Control for Unobservables." *Review of Economic Studies*,

70(2): 317–341.

Lindley, Joanne. 2009. "The Over-Education of UK Immigrants and Minority Ethnic Groups: Evidence from the Labour Force Survey." *Economics of Education Review*, 28(1): 80 – 89.

Mayr, Karin, and Giovanni Peri. 2009. "Brain Drain and Brain Return: Theory and Application to Eastern-Western Europe." *The B.E. Journal of Economic Analysis & Policy*, 9(1): 1–52.

Mitaritonna, Cristina, Gianluca Orefice, and Giovanni Peri. 2017. "Immigrants and firms' outcomes: Evidence from France." *European Economic Review*, 96: 62 – 82.

Mollisi, Vincenzo, and Gabriele Rovigatti. 2017. "Theory and practice of TFP estimation: The control function approach using stata." CEIS Working Paper 399.

Nielsen, Chantal. 2011. "Immigrant over-education: evidence from Denmark." *Journal of Population Economics*, 24(2): 499 – 520.

OECD. 2013. *Coping with Emigration in Baltic and East European Countries.* OECD Publishing, Paris.

Olley, G Steven, and Ariel Pakes. 1996. "The Dynamics of Productivity in the Telecommunications Equipment Industry." *Econometrica*, 64(6): 1263–97.

Ortega, Francesc, and Giovanni Peri. 2013. "The effect of income and immigration policies on international migration." *Migration Studies*, 1(1): 47–74.

Ottaviano, Gianmarco, and Giovanni Peri. 2013. "New Frontiers Of Immigration Research: Cities And Firms." *Journal of Regional Science*, 53(1): 1–7.

Ottaviano, Gianmarco I.P., Giovanni Peri, and Greg C. Wright. 2018. "Immigration, trade and productivity in services: Evidence from U.K. firms." *Journal of International Economics*, 112: 88–108.

Paserman, M Daniele. 2013. "Do high-skill immigrants raise productivity? Evidence from Israeli manufacturing firms, 1990-1999." *IZA Journal of Migration*, 2(1): 1–31.

Peri, Giovanni. 2012. "The Effect Of Immigration On Productivity: Evidence From U.S. States." *The Review of Economics and Statistics*, 94(1): 348–358.

Peri, Giovanni, Kevin Shih, and Chad Sparber. 2015. "STEM Workers, H-1B Visas, and Productivity in US Cities." *Journal of Labor Economics*, 33(S1): 225–255.

Rojas-Romagosa, Hugo, and Johannes Bollen. 2018. "Estimating migration from the EU's free movement of people principle." CPB Netherlands Bureau for Economic Policy Analysis.

Syverson, Chad. 2011. "What determines productivity?" *Journal of Economic literature*, 49(2): 326–65.

Topalova, Petia, and Amit Khandelwal. 2011. "Trade liberalization and firm pro-

ductivity: The case of India." *Review of Economics and Statistics*, 93(3): 995–1009.

Visintin, Stefano, Kea Tijdens, and Maarten van Klaveren. 2015. "Skill Mismatch among Migrant Workers: Evidence from a Large Multi-country Dataset." *IZA Journal of Migration*, 4.

Wooldridge, Jeffrey M. 2009. "On estimating firm-level production functions using proxy variables to control for unobservables." *Economics Letters*, 104(3): 112 – 114.

Zaiceva, Anzelika. 2014. "Post-enlargement emigration and new EU members' labor markets." *IZA World of Labor*.

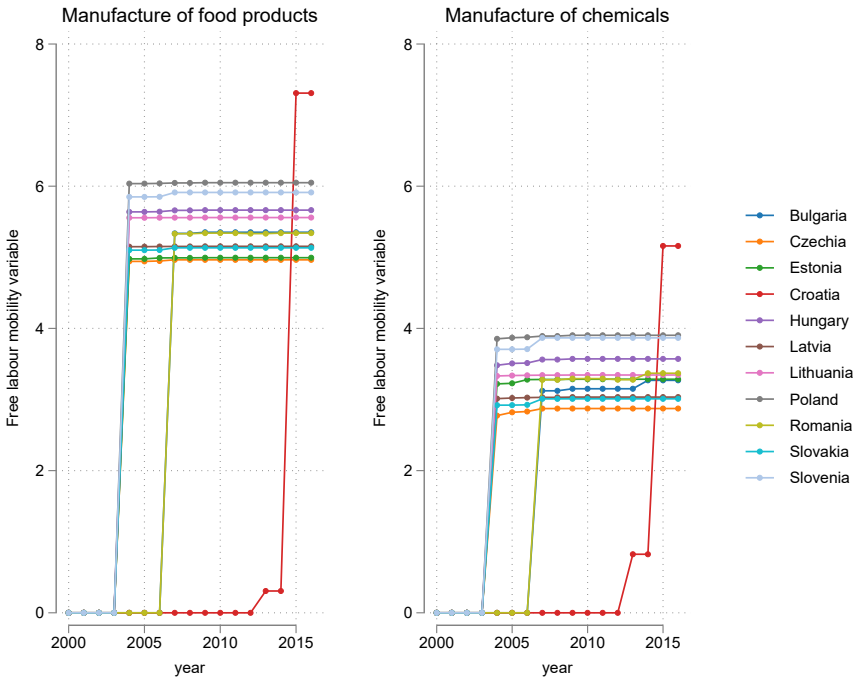
A Appendix (For Online Publication)

A.1 Additional Figures

Figure A1: **Variation in the Free Labor Mobility Variable**

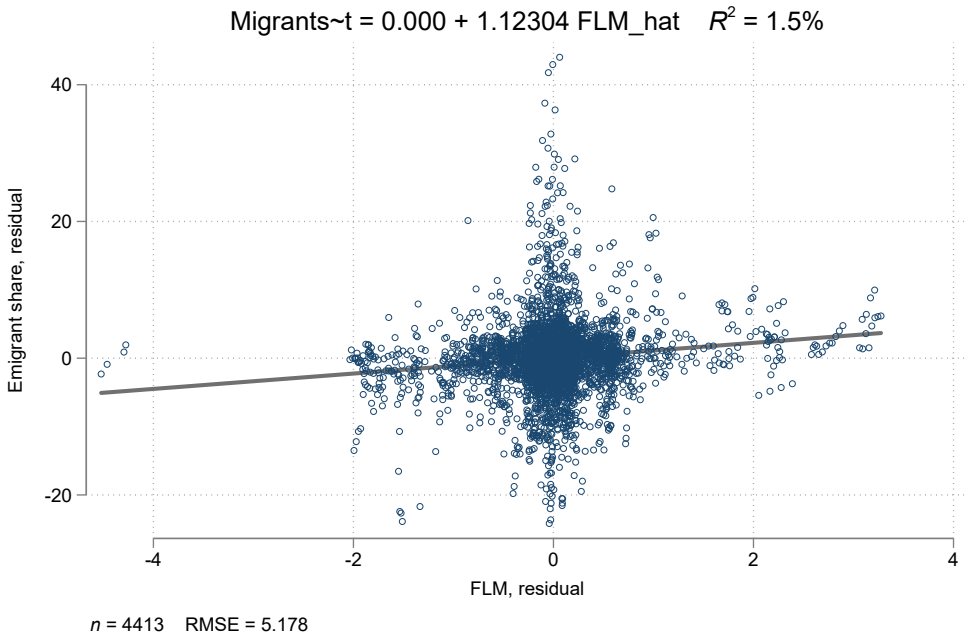
Notes: This graph shows the variation in the FLM variable. We compare different industries (y-Axis) in different countries (x-Axis) in 2004 and 2013. The darker the shading, the stronger a specific industry in a specific country in a given year is exposed to the emigration of workers. The shading reflects the magnitude of the FLM variable, which can be interpreted as the number (in natural logarithm + 1) of jobs available to NMS workers from a given origin and industry without any restrictions in EU15+4. One can see, for instance, that Croatia (HR), Bulgaria (BG) and Romania (RO) are completely in white in 2004 because those countries had not yet joined the EU and therefore emigration opportunities were restricted. In 2007, however, Romania and Bulgaria joined the EU, and certain sectors in certain destination countries already lifted labor-market restrictions for them, therefore creating better emigration opportunities.

Figure A2: **Illustration of Changes in the Free Labor Mobility Variable**



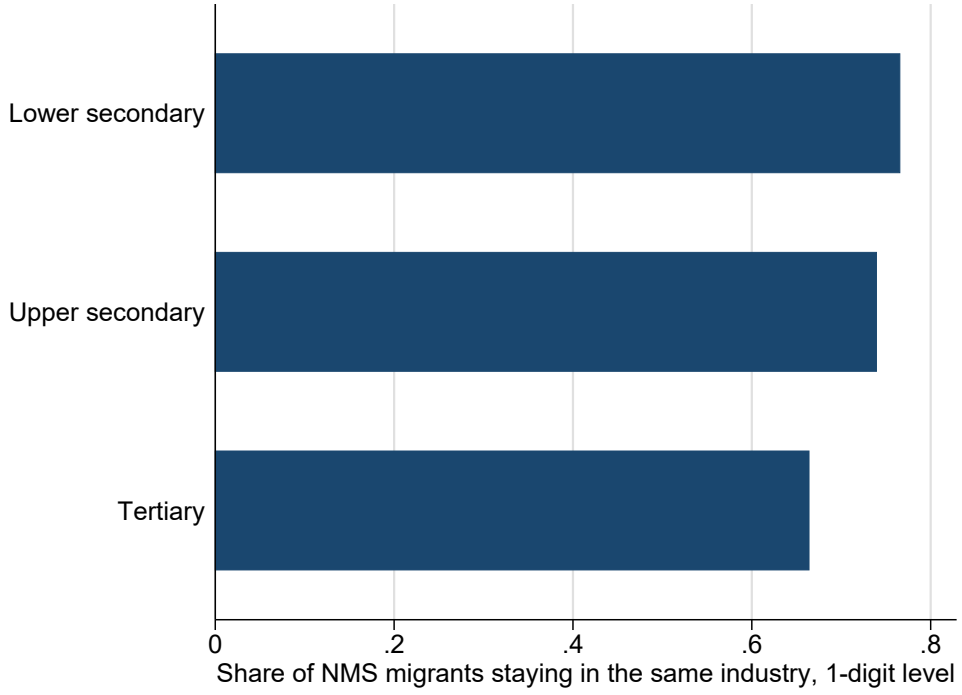
Notes: This graph illustrates variation in the FLM variable by showing an example of two industries: food manufacturing and production of chemical products. In total there are 24 different industries within the manufacturing sector. One can see that different countries and different sectors are affected differently due to variation in the time of sector opening, accession to the EU and due to differences in proximity.

Figure A3: **First Stage Illustration**



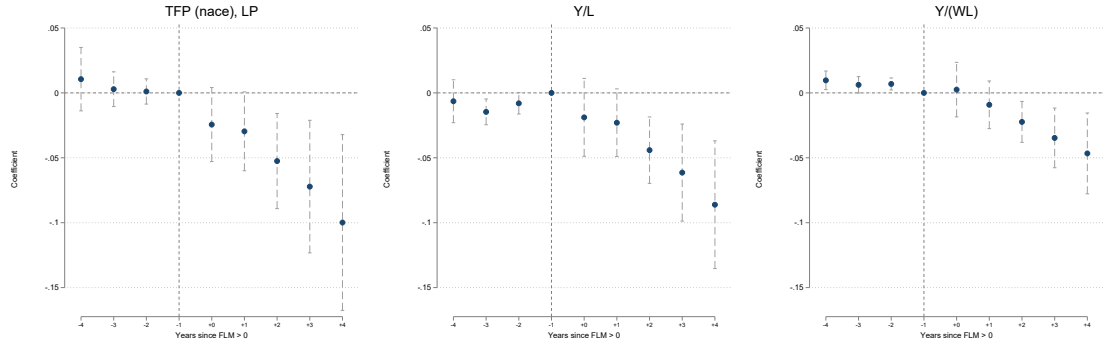
Notes: The scatter plot shows the correlation between the residuals of the emigrant share and the FLM variable after partialling out origin-industry, origin-year specific and industry-year specific fixed effects. In addition to the fixed effects, we control for differential time trends at the origin-industry level, by interacting average growth rate in employment, value-added, investment and import share between 2000 and 2003 with a linear time trend.

Figure A4: **Do NMS Migrants Stay in the Same Industry in Destination Countries?**



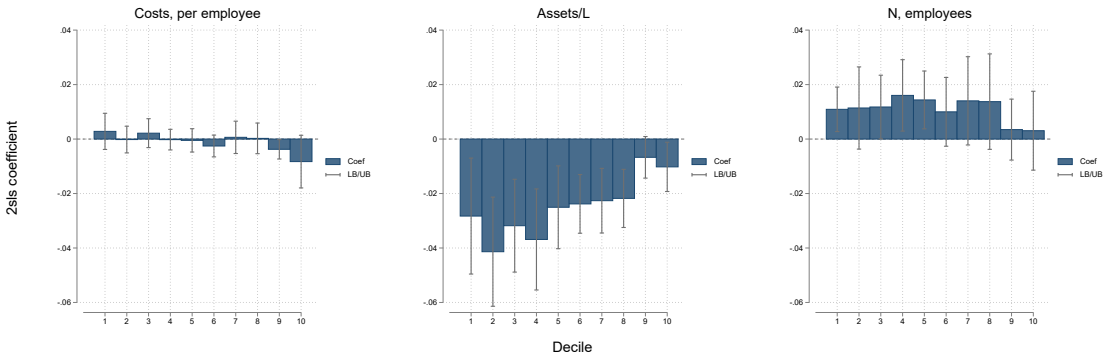
Source: EU LFS 2005-19. *Notes:* The sample includes migrants from NMS countries who are between 20 and 64 years old and who moved to another EU Member State in the year of the survey and who worked both in the previous and the current countries of residence. The graph indicates the share of all movers that report working in the same industry after their move.

Figure A5: **The Effect of Exposure to Emigration on Firm Productivity, Event-Study Specification**



Notes: This figure presents regression results for the event-study specification. The sample includes firms that experience the first labor-market opening (FLM changes from zero to a positive value) and that appear in the data for at least nine periods. We center the data around the first opening, hence period "zero" corresponds to the first year when FLM becomes positive for a given firm. The negative productivity effects take place with a delay because the regressor is the law change and not subsequent emigration. In the regressions we can exploit more fine-grained variation and can go beyond the aggregations of the event study. We refrain from introducing lags into regressions because it would be impossible to disentangle the lagged response of policy changes from subsequent policy changes. The figure shows the estimated coefficients of the interactions between time periods before and after the opening with a maximum FLM exposure that a given firm faces in our sample. Upper and lower bounds correspond to the 95 percent confidence interval. The regressions control for the firm, origin-year, and industry-year fixed effects, and for differential time trends at the origin-industry level. Additional controls include firm age and age squared. Standard errors are clustered at the origin-industry level.

Figure A6: **Other Firm-Level Outcomes Along the Productivity Distribution**



Notes: This figure presents regression results for other outcomes for incumbent firms along different productivity deciles. The sample only includes firms that existed prior to 2004 as firms are divided into productivity deciles based on their average productivity between 2000 and 2003. On the x-axes we indicate firm productivity deciles and on the y-axes we indicate the estimated coefficients from a 2SLS regression with the respective outcome indicated at the top of each graph. The blue bars show the estimated coefficients for instrumented migration and the black lines show the upper and lower bounds of the 95 percent confidence interval.

A.2 Additional Tables

Table A1: **Overview of the Gradual Opening of the EU15+4 Labor Markets**

Country	NMS10 (2004 entry)	NMS3 (2007 entry: Bulgaria, Romania)	NMS3 (2013 entry: Croatia)	Sectoral Exceptions
Austria	2011	2014	2020	NMS10 (2007-2010), NMS3 (2007-2013): Construction, Manufacturing of Electronics and Metals, Food and beverage services (restaurant business), other sectors with labor shortages
Belgium	2009	2014	2015	-
Denmark	2009	2009	2013	-
Finland	2006	2007	2013	-
France	2008	2014	2015	NMS10 (2005-2007), NMS3 (2007-2013): Agriculture, Construction, Accommodation and food services (tourism and catering), other sectors with labor shortages
Germany	2011	2014	2015	NMS10 (2004-2010), NMS3 (2007-2013): sectors with labor shortages
Greece	2006	2009	2015	-
Iceland	2006	2012	2015	-
Ireland	2004	2012	2013	-
Italy	2006	2012	2015	NMS10 (2004-2005): sectors with labor shortages; NMS3 (2007-2011): Agriculture, Construction, Engineering, Accommodation and food services (tourism and catering), Domestic work and care services, other sectors with labor shortages; Occupations: Managerial and professional occupations
Luxembourg	2008	2014	2015	NMS3 (2007 - 2013): Agriculture, Viticulture, Accommodation and food services (tourism and catering)
Netherlands	2007	2014	2018	NMS10 (2004-2006), NMS3 (2007-2013): International transport, Inland shipping, Health, Slaugther-house/meat-packaging, other sectors with labor shortages
Norway	2009	2012	2014	NMS10 (2004-2008), NMS3 (2007-2011): sectors with labor shortages
Portugal	2006	2009	2013	-
Spain	2006	2009	2015	Reintroduction of restrictions for Romanians: 11/08/2011 - 31/12/2013
Sweden	2004	2007	2013	-
Switzerland	2011	2016	2024 (tbc)	-
United Kingdom	2004	2014	2018	NMS3 (2007-2013): Agriculture, Food manufacturing

Notes: Column 2 shows the year of the labor market opening of the respective country for NMS10 countries, Column 3 shows the year of the labor market opening of the respective country for Bulgaria and Romania and Column 4 for Croatia. Column 5 shows, which industries were exempt from restrictions before the transitional provisions for the entire labor market.
Source: Compiled by the authors using the LABREF database (European Commission) and national legislations.

Table A2: **Summary Statistics, Industry-Level Variables**

	Mean	SD	Min	Max	N
Emigration share, % (oit)	9.664	12.404	.0144	89.885	4413
Emigrants, stock (oit)	2802.327	6447.286	4.54	1.06e+05	4488
Free labour mobility (oit)	2.789	2.304	0	8.598	4488
Value added, mio EUR (oit)	542.329	934.398	-319	8723.800	3869
Number of employees (oit)	29209.009	43544.070	.434	3.91e+05	4488
Investment, mio EUR (oit)	134.341	225.451	.0000394	2177.700	4488
Import share in industry turnover, % (oit)	32.666	33.860	.0128	100.000	4488
Skill shortages in EU19, share of firms, % (it)	5.004	3.592	.0438	24.037	4488
Skill shortages in NMS, share of firms, % (oit)	13.013	13.805	0	99.450	3350
High product-market competition, share of firms, % (oit)	10.241	10.161	0	94.750	2693
Had training, share of workers, % (oit)	0.869	2.527	0	23.762	3188
Job-related training, share of workers, % (oit)	0.323	1.538	0	23.762	3188

Notes: The table presents summary statistics for the variables at the industry level. We show the level of variation of all variables in parentheses: o-origin, i-industry, t-year. The main independent variable is the emigration rate (in percent), which we calculate as the ratio between the current stock of emigrants and the industry's average employment in 2000-2002. The median emigration rate in the sample is 5.19 percent. The maximum emigration rate in a given country, industry, year seems to be very high (89.89 percent). We investigated this further, and it concerns mainly sector 33 (Repair and installation of machinery and equipment) in Romania and Bulgaria after the accession in 2007. Our results hold when excluding Romania (Table A13) and reestimating our baseline specification without the five percent of observations that have the highest emigration rate (Table A14) as shown in the robustness checks.

Table A3: **First Stage Regressions with Different Instruments, Full Sample**

	(1)	(2)	(3)	(4)
	Legal dummies&Market size I	Legal dummies&Market size I	Legal dummies&Market size II	Legal dummies&Market size II
VARIABLES	Distance	Diaspora	Distance	Diaspora
FLM_{oit}	0.896	1.125	0.776	0.998
	(0.256)	(0.286)	(0.243)	(0.354)
Observations	4,413	4,413	4,413	4,413
R-squared	0.825	0.826	0.825	0.825
Clusters	263	263	263	263
F-stat	12.26	15.47	10.24	7.940
Mean emigration share (percent)	9.664	9.664	9.664	9.664
SD emigration share	12.40	12.40	12.40	12.40
Mean Δ emigration share	0.682	0.682	0.682	0.682
SD Δ emigration share	5.541	5.541	5.541	5.541
Origin#Industry FE	✓	✓	✓	✓
Year#Origin FE	✓	✓	✓	✓
Year#Industry FE	✓	✓	✓	✓
Linear trends	✓	✓	✓	✓
Level	Industry	Industry	Industry	Industry

Notes: The table presents different versions of the first stage: the effect of higher free labor mobility (FLM) on emigration from NMS. The sample includes all observations with available data on emigration, FLM, and controls. The dependent variable is $\frac{Emigrants_{oit}}{Employees_{oi,pre}} * 100$, where the numerator is the stock of emigrants from a given origin, working in a given industry and year in one of the EU15+4 countries and the denominator is the average number of employees in a given origin country and industry before 2004. The Columns compare first stage results for different variations of the FLM variable. As proximity weights, we use either inverse bilateral distances (Columns 1 and 3) or previous bilateral distribution of emigrants (Columns 2 and 4). Bilateral distances are taken from Conte, Cotterlaz and Mayer (2022), who "construct country-to-country distances that take into account the spatial distribution of economic activity within each country". Market size represents the *pre-determined* capacity of destination industries to absorb migrants from NMS. Market size I is the number of employees multiplied by the share of firms reporting labor shortages in a given destination industry as of 2000-2003. Market size II is the number of employees multiplied by the share of employed migrants in a given destination industry. All specifications control for origin-industry, origin-year, and industry-year fixed effects. In addition to the fixed effects, we control for differential time trends at the origin-industry level, by interacting average growth rate in employment, value-added, investment and import share between 2000 and 2003 with a linear time trend. Standard errors (in parentheses) are clustered at the origin-industry level.

Table A4: **Pre-Trend Analysis: Do Future Changes in FLM Predict Changes in Outcomes Before 2004**

	(1)	(2)	(3)	(4)
	Δ Emigration	Δ Y/L	Δ Exits	Δ N firms
VARIABLES	2000-2003	2000-2003	2000-2003	2000-2003
FLM_{oi2004}	0.006	-0.018	0.000	0.015
	(0.007)	(0.013)	(0.013)	(0.016)
FLM_{oi2007}	0.009	-0.015	0.016	0.010
	(0.007)	(0.015)	(0.017)	(0.012)
FLM_{oi2013}	0.010	-0.015	0.016	0.010
	(0.007)	(0.015)	(0.016)	(0.012)
Observations	264	206	225	218
R-squared	0.578	0.312	0.689	0.635
Clusters	264	206	225	218
F-stat	2.469	1.062	1	0.723
Origin FE	✓	✓	✓	✓
Industry FE	✓	✓	✓	✓
Level	Industry	Industry	Industry	Industry

Notes: Every Column presents the results of regressions of an industry-level outcome (average annual change in 2000-2003) on future changes in the FLM variable. We separately estimate the effects of changes in the FLM for 2004-2003, 2007-2003, and 2013-2003, because 2004, 2007 and 2013 correspond to a new wave of the EU enlargement. The FLM corresponds to the FLM used in our baseline estimations (Column 2 in Table A3). Standard errors (in parentheses) are clustered at the origin-industry level. All specifications comprise origin and industry fixed effects. Note that the number of observations varies due to differences in data availability. The results hold when we restrict the sample to industries with no missing data for all four outcomes.

Table A5: **Firm Productivity and Free Labor Mobility: 10-Year Differences**

VARIABLES	(1) 2SLS TFP (nace) 2013-2003	(2) 2SLS Y/L 2013-2003	(3) 2SLS Y/(WL) 2013-2003
Emigration, 2013-2003	-0.046	-0.037	-0.016
	(0.013)	(0.011)	(0.005)
Observations	32,062	32,062	32,062
Clusters	198	198	198
Mean outcome 2003	4.965	8.602	1.489
SD outcome 2003	1.471	1.462	0.877
Mean Δ emigration share 2013-2003	10.28	10.28	10.28
SD Δ emigration	15.09	15.09	15.09
First stage F-stat	10.25	10.25	10.25
FLM coefficient	3.848	3.848	3.848
FLM se	1.202	1.202	1.202
Origin	✓	✓	✓
Industry	✓	✓	✓
Controls	✓	✓	✓

Notes: The table presents results from 2SLS regressions of the 10-year within-firm difference in our three productivity measures on the 10-year difference in the emigration rate. *Median change* in the emigration rate between 2013 and 2003: 3.589 percentage points. The sample includes firms that we observe in the sample in both 2003 and 2013. All specifications are estimated with origin country and industry fixed effects. Additional controls include age and age squared, as well as average growth rate in employment, value-added, investment and import share between 2000 and 2003. Standard errors (in parentheses) are clustered at the origin-industry-year level.

Table A6: **Reduced Form Results for Firm Productivity: Leads and Lags of the Free Labor Mobility Variable**

VARIABLES	(1) Reduced form TFP (nace)	(2) Reduced form Y/L	(3) Reduced form Y/(WL)
FLM_{oit+2}	0.000	0.010	0.004
	(0.008)	(0.006)	(0.004)
FLM_{oit+1}	-0.006	-0.003	-0.004
	(0.006)	(0.005)	(0.003)
FLM_{oit}	-0.003	0.008	0.003
	(0.020)	(0.022)	(0.017)
FLM_{oit-1}	-0.017	-0.015	-0.009
	(0.018)	(0.016)	(0.011)
FLM_{oit-2}	-0.075	-0.069	-0.038
	(0.031)	(0.031)	(0.021)
Observations	625,467	625,467	625,467
R-squared	0.802	0.744	0.732
Clusters	238	238	238
Mean outcome pre-2004	4.958	8.448	1.373
SD outcome pre-2004	1.325	1.210	0.610
Firm	✓	✓	✓
Year#Origin	✓	✓	✓
Year#Industry	✓	✓	✓
Linear trends/controls	✓	✓	✓

Notes: The table presents the reduced form regressions that include two lags, the contemporaneous value and two forwards of the FLM variable. The sample includes firms that we observe in the sample for at least five periods. In addition to the fixed effects mentioned in the table, we control for differential time trends at the origin-industry level by interacting average growth rate in employment, value-added, investment and import share between 2000 and 2003 with a linear time trend. Additional controls include firm age and age squared. Standard errors (in parentheses) are clustered at the origin-industry level.

Table A7: **The Effect of Emigration on Firm Outcomes, Incumbent Firms**

VARIABLES	(1) 2SLS TFP (nace) LP	(2) 2SLS Y/L	(3) 2SLS Y/(WL)
$Emigration_{oit}$	-0.025	-0.018	-0.012
	(0.007)	(0.005)	(0.004)
Observations	744,625	744,625	744,625
Clusters	224	224	224
Mean outcome pre-2004	4.897	8.401	1.443
SD outcome pre-2004	1.496	1.457	0.786
Mean Δ emigration share	1.069	1.069	1.069
SD Δ emigration	6.822	6.822	6.822
First stage F-stat	14.02	14.02	14.02
FLM coefficient	2.874	2.874	2.874
FLM se	0.768	0.768	0.768
Firm	✓	✓	✓
Year#Origin	✓	✓	✓
Year#Industry	✓	✓	✓
Linear trends/controls	✓	✓	✓

Notes: The table presents 2SLS effects of emigration on three measures of firm productivity for incumbent firms (those that were already in the sample prior to 2004). Column 1 shows the effect of emigration on TFP (estimated according to the Levinson Petrin method). Column 2 shows the result for labor productivity (value-added/employees) and Column 3 for wage-adjusted labor productivity (value-added/employee costs). In addition to the fixed effects mentioned in the table, we control for differential time trends at the origin-industry level, by interacting average growth rate in employment, value-added, investment and import share between 2000 and 2003 with a linear time trend. Additional controls include firm age and age squared. Standard errors (in parentheses) are clustered at the origin-industry level.

Table A8: **Industry-level Effects, Eurostat Data (Only Industries with Good Orbis Coverage)**

	(1) 2SLS Exits	(2) 2SLS N firms	(3) 2SLS Y/L	(4) 2SLS Y/(WL)	(5) 2SLS Mean TFP	(6) 2SLS Mean TFP new firms
VARIABLES	Eurostat	Eurostat	Eurostat	Eurostat	Orbis	Orbis
$Emigration_{oit}$	0.109	0.003	0.020	0.028	0.004	0.097
	(0.056)	(0.015)	(0.031)	(0.024)	(0.012)	(0.041)
Observations	2,937	2,869	2,720	2,703	3,159	3,159
Clusters	192	189	185	185	192	192
Mean outcome pre-2004	2.968	6.393	2.286	0.988	5.504	1.551
SD outcome pre-2004	2.270	1.915	0.688	0.272	0.970	1.563
Mean Δ emigration share	0.540	0.540	0.540	0.540	0.540	0.540
SD Δ emigration	4.502	4.502	4.502	4.502	4.502	4.502
First stage F-stat	11.03	12.82	12.96	13.06	11.81	11.81
FLM coefficient	1.168	1.292	1.428	1.462	1.200	1.200
FLM se	0.352	0.361	0.397	0.404	0.349	0.349
Origin#Industry FE	✓	✓	✓	✓	✓	✓
Year#Origin FE	✓	✓	✓	✓	✓	✓
Year#Industry FE	✓	✓	✓	✓	✓	✓
Linear trends/controls	✓	✓	✓	✓	✓	✓

Notes: The table presents 2SLS regressions of industry-level outcomes on emigration. The sample is limited to origin-industry-year observations with good Orbis coverage (without Hungary, Lithuania, and Latvia). Column 1 shows the number of firm exits and Column 2 the number of firms. Columns 3 and 4 use Eurostat data: Y/L is calculated by dividing the total value-added on the total number of workers, $Y/(WL)$ is calculated by dividing the total value-added on the total employee costs. Column 5 uses Orbis data to calculate the average TFP of all present firms in a given origin-industry-year. Column 6 uses Orbis data to calculate the average first-year TFP of entering firms. In addition to the fixed effects mentioned in the table, we control for differential time trends at the origin-industry level, by interacting average growth rate in employment, value-added, investment and import share between 2000 and 2003 with a linear time trend. Standard errors (in parentheses) are clustered at the origin-industry level. Note that the number of observations and hence the first-stage statistics in Columns 1-6 vary due to differences in data availability. The results hold when we, in addition, restrict the sample to industries with no missing data for all six outcomes.

Table A9: **Industry-Level Outcomes: 10-Year Differences**

	(1) 2SLS N firms	(2) 2SLS Y/L	(3) 2SLS Y/(WL)	(4) 2SLS Mean TFP	(5) 2SLS Mean TFP (size-weighted)	(6) 2SLS Mean TFP (size-weighted) always present
VARIABLES	Eurostat 2013-2003	Eurostat 2013-2003	Eurostat 2013-2003	Orbis 2013-2003	Orbis 2013-2003	Orbis 2013-2003
Emigration, 2013-2003	-0.026	0.033	0.035	0.060	0.088	-0.003
	(0.024)	(0.041)	(0.029)	(0.037)	(0.045)	(0.037)
Observations	214	198	197	206	206	206
Clusters	214	198	197	206	206	206
Mean outcome 2003	6.231	2.440	1.050	5.542	5.971	5.752
SD outcome 2003	1.950	0.690	0.256	0.937	0.982	1.502
Mean Δ emigration share 2013-2003	5.247	5.247	5.247	5.247	5.247	5.247
SD Δ emigration	11.13	11.13	11.13	11.13	11.13	11.13
First stage F-stat	12.84	9.468	9.548	6.884	6.884	6.884
FLM coefficient	1.595	1.528	1.538	1.244	1.244	1.244
FLM se	0.445	0.497	0.498	0.474	0.474	0.474
Origin	✓	✓	✓	✓	✓	✓
Industry	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓

Notes: The table presents results from 2SLS regressions of 10-year differences in the different industry-level outcomes on 10-year differences in the emigration rate. The main outcomes are the number of firms (Column 1), labor productivity (Column 2), wage-adjusted labor productivity (Column 3), the average TFP (Column 4), the average *size-weighted* TFP (Column 5), and the average *size-weighted* TFP of always-present firms between 2003 and 2013 (Column 6) in a given industry, origin and year cell. We use Eurostat data to obtain outcomes in Columns 1-3 and Orbis data for results in Columns 4-6. In addition to the fixed effects mentioned in the table, we control for the average growth rate in employment, value-added, investment and import share between 2000 and 2003. Standard errors (in parentheses) are clustered at the origin-industry level. Note that the number of observations and hence the first-stage statistics in Columns 1-6 vary due to differences in data availability. The results hold when we restrict the sample to industries with no missing data for all six outcomes.

Table A10: **Additional Firm Outcomes and Free Labor Mobility: 10-Year Differences**

VARIABLES	(1) 2SLS Costs per employee 2013-2003	(2) 2SLS Assets/L 2013-2003	(3) 2SLS N employees 2013-2003
Emigration, 2013-2003	-0.013	-0.035	0.007
	(0.005)	(0.009)	(0.007)
Observations	32,062	32,062	32,062
Clusters	198	198	198
Mean outcome 2003	7.511	9.173	2.493
SD outcome 2003	1.303	1.512	1.430
Mean Δ emigration share 2013-2003	10.28	10.28	10.28
SD Δ emigration	15.09	15.09	15.09
First stage F-stat	10.25	10.25	10.25
FLM coefficient	3.848	3.848	3.848
FLM se	1.202	1.202	1.202
Origin	✓	✓	✓
Industry	✓	✓	✓
Controls	✓	✓	✓

Notes: The table presents results from 2SLS regressions of the 10-year within-firm difference in three additional outcomes on the 10-year difference in the emigration rate. *Median change* in the emigration rate between 2013 and 2003: 3.589 percentage points. The sample includes firms that we observe in the sample in both 2003 and 2013. All specifications are estimated with origin country and industry fixed effects. Additional controls include age and age squared, as well as average growth rate in employment, value-added, investment and import share between 2000 and 2003. Standard errors (in parentheses) are clustered at the origin-industry-year level.

Table A11: **Regression Restricted to Firms with 5-250 Employees**

VARIABLES	(1) 2SLS TFP (nace) LP	(2) 2SLS Y/L	(3) 2SLS Y/(WL)	(4) 2SLS Costs per employee	(5) 2SLS Assets/L	(6) 2SLS N employees
$Emigration_{oit}$	-0.019	-0.013	-0.011	0.002	-0.019	0.014
	(0.007)	(0.006)	(0.004)	(0.002)	(0.005)	(0.006)
Observations	957,917	957,917	957,917	957,917	957,917	957,917
Clusters	238	238	238	238	238	238
Mean outcome pre-2004	5.097	8.487	1.365	7.515	8.988	2.985
SD outcome pre-2004	1.389	1.330	0.612	1.044	1.398	1.065
Mean Δ emigration share	0.999	0.999	0.999	0.999	0.999	0.999
SD Δ emigration	6.507	6.507	6.507	6.507	6.507	6.507
First stage F-stat	14.77	14.77	14.77	14.77	14.77	14.77
FLM coefficient	2.636	2.636	2.636	2.636	2.636	2.636
FLM se	0.686	0.686	0.686	0.686	0.686	0.686
Firm FE	✓	✓	✓	✓	✓	✓
Year#Origin FE	✓	✓	✓	✓	✓	✓
Year#Industry FE	✓	✓	✓	✓	✓	✓
Linear trends/controls	✓	✓	✓	✓	✓	✓

Notes: The table presents 2SLS effects of emigration on various firm outcomes. The sample excludes very small firms (less than 5 employees) and very large firms (more than 250 employees). The Columns show the results for the following outcomes: firm TFP, labor productivity, wage-adjusted labor productivity, costs per employee, capital-labor ratio and the number of employees. In addition to the fixed effects mentioned in the table, we control for differential time trends at the origin-industry level, by interacting average growth rate in employment, value-added, investment and import share between 2000 and 2003 with a linear time trend. Additional controls include firm age and age squared. Standard errors (in parentheses) are clustered at the origin-industry level.

Table A12: **Regression Excluding Years of the Financial Crisis 2009-11**

VARIABLES	(1) 2SLS TFP (nace) LP	(2) 2SLS Y/L	(3) 2SLS Y/(WL)	(4) 2SLS Costs per employee	(5) 2SLS Assets/L	(6) 2SLS N employees
$Emigration_{oit}$	-0.024	-0.014	-0.010	-0.001	-0.025	0.009
	(0.008)	(0.006)	(0.005)	(0.003)	(0.007)	(0.006)
Observations	1,163,714	1,163,714	1,163,714	1,163,714	1,163,714	1,163,714
Clusters	239	239	239	239	239	239
Mean outcome pre-2004	4.899	8.401	1.443	7.360	8.945	2.533
SD outcome pre-2004	1.496	1.457	0.786	1.168	1.517	1.424
Mean Δ emigration share	1.356	1.356	1.356	1.356	1.356	1.356
SD Δ emigration	6.245	6.245	6.245	6.245	6.245	6.245
First stage F-stat	13.36	13.36	13.36	13.36	13.36	13.36
FLM coefficient	2.371	2.371	2.371	2.371	2.371	2.371
FLM se	0.649	0.649	0.649	0.649	0.649	0.649
Firm FE	✓	✓	✓	✓	✓	✓
Year#Origin FE	✓	✓	✓	✓	✓	✓
Year#Industry FE	✓	✓	✓	✓	✓	✓
Linear trends/controls	✓	✓	✓	✓	✓	✓

Notes: The table presents 2SLS effects of emigration on various firm outcomes. The sample excludes years of the financial crisis 2009-11. The Columns show the results for the following outcomes: firm TFP, labor productivity, wage-adjusted labor productivity, costs per employee, capital-labor ratio and the number of employees. In addition to the fixed effects mentioned in the table, we control for differential time trends at the origin-industry level, by interacting average growth rate in employment, value-added, investment and import share between 2000 and 2003 with a linear time trend. Additional controls include firm age and age squared. Standard errors (in parentheses) are clustered at the origin-industry level.

Table A13: **Regression Excluding Romania, Hungary, Lithuania and Latvia**

VARIABLES	(1) 2SLS TFP (nace) LP	(2) 2SLS Y/L	(3) 2SLS Y/(WL)	(4) 2SLS Costs per employee	(5) 2SLS Assets/L	(6) 2SLS N employees
$Emigration_{oit}$	-0.046	-0.048	-0.038	0.006	-0.031	0.044
	(0.024)	(0.023)	(0.018)	(0.005)	(0.017)	(0.020)
Observations	783,077	783,077	783,077	783,077	783,077	783,077
Clusters	168	168	168	168	168	168
Mean outcome pre-2004	5.414	9.292	1.391	8.284	9.814	2.920
SD outcome pre-2004	1.318	1.208	0.679	0.987	1.276	1.444
Mean Δ emigration share	0.623	0.623	0.623	0.623	0.623	0.623
SD Δ emigration	3.533	3.533	3.533	3.533	3.533	3.533
First stage F-stat	8.198	8.198	8.198	8.198	8.198	8.198
FLM coefficient	0.705	0.705	0.705	0.705	0.705	0.705
FLM se	0.246	0.246	0.246	0.246	0.246	0.246
Firm FE	✓	✓	✓	✓	✓	✓
Year#Origin FE	✓	✓	✓	✓	✓	✓
Year#Industry FE	✓	✓	✓	✓	✓	✓
Linear trends/controls	✓	✓	✓	✓	✓	✓

Notes: The table presents 2SLS effects of emigration on various firm outcomes. The sample excludes Romania (46,6 percent of all observations) and Hungary, Lithuania, and Latvia (poor Orbis coverage). The Columns show the results for the following outcomes: firm TFP, labor productivity, wage-adjusted labor productivity, costs per employee, capital-labor ratio and the number of employees. In addition to the fixed effects mentioned in the table, we control for differential time trends at the origin-industry level, by interacting average growth rate in employment, value-added, investment and import share between 2000 and 2003 with a linear time trend. Additional controls include firm age and age squared. Standard errors (in parentheses) are clustered at the origin-industry level.

Table A14: **Regression Excluding Five Percent of Observations with the Highest Emigration Rate**

VARIABLES	(1) 2SLS TFP (nace) LP	(2) 2SLS Y/L	(3) 2SLS Y/(WL)	(4) 2SLS Costs per employee	(5) 2SLS Assets/L	(6) 2SLS N employees
$Emigration_{oit}$	-0.028	-0.018	-0.013	-0.000	-0.024	0.010
	(0.008)	(0.006)	(0.004)	(0.003)	(0.004)	(0.005)
Observations	1,401,919	1,401,919	1,401,919	1,401,919	1,401,919	1,401,919
Clusters	237	237	237	237	237	237
Mean outcome pre-2004	4.899	8.400	1.441	7.361	8.946	2.535
SD outcome pre-2004	1.497	1.457	0.785	1.169	1.517	1.425
Mean Δ emigration share	0.812	0.812	0.812	0.812	0.812	0.812
SD Δ emigration	5.672	5.672	5.672	5.672	5.672	5.672
First stage F-stat	17.69	17.69	17.69	17.69	17.69	17.69
FLM coefficient	2.295	2.295	2.295	2.295	2.295	2.295
FLM se	0.546	0.546	0.546	0.546	0.546	0.546
Firm	✓	✓	✓	✓	✓	✓
Year#Origin	✓	✓	✓	✓	✓	✓
Year#Industry	✓	✓	✓	✓	✓	✓
Linear trends/controls	✓	✓	✓	✓	✓	✓

Notes: The table presents 2SLS effects of emigration on various firm outcomes. The sample excludes five percent of observations with the highest emigration rate. The Columns show the results for the following outcomes: firm TFP, labor productivity, wage-adjusted labor productivity, costs per employee, capital-labor ratio and the number of employees. In addition to the fixed effects mentioned in the table, we control for differential time trends at the origin-industry level, by interacting average growth rate in employment, value-added, investment and import share between 2000 and 2003 with a linear time trend. Additional controls include firm age and age squared. Standard errors (in parentheses) are clustered at the origin-industry level.

Table A15: **Robustness to Different Measures of TFP**

	(1)	(2)	(3)
	2SLS	2SLS	2SLS
	TFP (nace)	TFP (nace)	TFP (all)
VARIABLES	LP	WDG	LP
$Emigration_{oit}$	-0.023***	-0.025***	-0.010*
	(0.007)	(0.007)	(0.006)
Observations	1,473,412	1,473,412	1,473,412
Clusters	239	239	239
Mean outcome pre-2004	4.899	4.582	4.818
SD outcome pre-2004	1.496	1.510	1.300
Mean Δ emigration share	1.023	1.023	1.023
SD Δ emigration	6.860	6.860	6.860
First stage F-stat	15.54	15.54	15.54
FLM coefficient	2.746	2.746	2.746
FLM se	0.697	0.697	0.697
Firm FE	✓	✓	✓
Year#Origin FE	✓	✓	✓
Year#Industry FE	✓	✓	✓
Linear trends/controls	✓	✓	✓

Notes: The table presents our baseline result in Column 1 and an alternative measure to calculate TPF (Wooldridge method) in Column 2. Column 3 shows TFP when calculated using Levinsohn-Petrin method but by assuming the same production function for all firms in the sample. In addition to the fixed effects mentioned in the table, we control for differential time trends at the origin-industry level, by interacting average growth rate in employment, value-added, investment and import share between 2000 and 2003 with a linear time trend. Additional controls include firm age and age squared. Standard errors (in parentheses) are clustered at the origin-industry level.

Table A16: **With Firm-Level Time Trends (Only Incumbents)**

VARIABLES	(1) 2SLS TFP (nace) LP	(2) 2SLS Y/L	(3) 2SLS Y/(WL)	(4) 2SLS Costs per employee	(5) 2SLS Assets/L	(6) 2SLS N employees
$Emigration_{oit}$	-0.023	-0.016	-0.012	-0.001	-0.023	0.011
	(0.007)	(0.006)	(0.004)	(0.002)	(0.005)	(0.006)
Observations	587,897	587,897	587,897	587,897	587,897	587,897
Clusters	204	204	204	204	204	204
Mean outcome pre-2004	4.899	8.401	1.443	7.360	8.945	2.533
SD outcome pre-2004	1.496	1.457	0.786	1.168	1.517	1.424
Mean Δ emigration share	1.069	1.069	1.069	1.069	1.069	1.069
SD Δ emigration	6.822	6.822	6.822	6.822	6.822	6.822
First stage F-stat	13.24	13.24	13.24	13.24	13.24	13.24
FLM coefficient	2.850	2.850	2.850	2.850	2.850	2.850
FLM se	0.783	0.783	0.783	0.783	0.783	0.783
Firm FE	✓	✓	✓	✓	✓	✓
Year#Origin FE	✓	✓	✓	✓	✓	✓
Year#Industry FE	✓	✓	✓	✓	✓	✓
Linear trends/controls	✓	✓	✓	✓	✓	✓

Notes: The table presents 2SLS regressions of emigration on three measures of firm productivity and three additional firm outcomes. In comparison to our baseline, this table includes firm-level linear trends. This table is restricted to incumbents for whom we can calculate the baseline for the linear trend as of 2000-2003. In addition to the fixed effects mentioned in the table, we control for differential time trends at the origin-industry level, by interacting average growth rate in employment, value-added, investment and import share between 2000 and 2003 with a linear time trend. Additional controls include firm age and age squared. Standard errors (in parentheses) are clustered at the origin-industry level.

Table A17: **Regression without Controls**

VARIABLES	(1) 2SLS TFP (nace) LP	(2) 2SLS Y/L	(3) 2SLS Y/(WL)	(4) 2SLS Costs per employee	(5) 2SLS Assets/L	(6) 2SLS N employees
$Emigration_{oit}$	-0.023	-0.015	-0.010	-0.002	-0.021	0.006
	(0.007)	(0.006)	(0.004)	(0.002)	(0.005)	(0.004)
Observations	1,473,412	1,473,412	1,473,412	1,473,412	1,473,412	1,473,412
Clusters	239	239	239	239	239	239
Mean outcome pre-2004	4.899	8.401	1.443	7.360	8.945	2.533
SD outcome pre-2004	1.496	1.457	0.786	1.168	1.517	1.424
Mean Δ emigration share	1.023	1.023	1.023	1.023	1.023	1.023
SD Δ emigration	6.860	6.860	6.860	6.860	6.860	6.860
First stage F-stat	18.59	18.59	18.59	18.59	18.59	18.59
FLM coefficient	3.006	3.006	3.006	3.006	3.006	3.006
FLM se	0.697	0.697	0.697	0.697	0.697	0.697
Firm FE	✓	✓	✓	✓	✓	✓
Year#Origin FE	✓	✓	✓	✓	✓	✓
Year#Industry FE	✓	✓	✓	✓	✓	✓
Linear trends/controls	X	X	X	X	X	X

Notes: The table presents 2SLS regressions of emigration on three measures of firm productivity and three additional firm outcomes. In comparison to our baseline, this table drops all control variables. The fixed effects are kept as shown in the table lines. Standard errors (in parentheses) are clustered at the origin-industry level.

Table A18: **Regression with an IV Generated Using a Random Forest Method**

VARIABLES	(1) 2SLS TFP (nace) LP	(2) 2SLS Y/L	(3) 2SLS Y/(WL)	(4) 2SLS Costs per employee	(5) 2SLS Assets/L	(6) 2SLS N employees
$Emigration_{oit}$	-0.022	-0.014	-0.007	-0.005	-0.027	0.005
	(0.008)	(0.006)	(0.004)	(0.002)	(0.005)	(0.005)
Observations	1,473,412	1,473,412	1,473,412	1,473,412	1,473,412	1,473,412
Clusters	239	239	239	239	239	239
Mean outcome pre-2004	4.899	8.401	1.443	7.360	8.945	2.533
SD outcome pre-2004	1.496	1.457	0.786	1.168	1.517	1.424
Mean Δ emigration share	1.023	1.023	1.023	1.023	1.023	1.023
SD Δ emigration	6.860	6.860	6.860	6.860	6.860	6.860
First stage F-stat	10.54	10.54	10.54	10.54	10.54	10.54
FLM coefficient	8.686	8.686	8.686	8.686	8.686	8.686
FLM se	2.676	2.676	2.676	2.676	2.676	2.676
Firm FE	✓	✓	✓	✓	✓	✓
Year#Origin FE	✓	✓	✓	✓	✓	✓
Year#Industry FE	✓	✓	✓	✓	✓	✓
Linear trends/controls	✓	✓	✓	✓	✓	✓

Notes: The table presents 2SLS regressions of emigration on three measures of firm productivity and three additional firm outcomes. In comparison to our baseline, this table uses a random forest method to create the instrument. In addition to the fixed effects mentioned in the table, we control for differential time trends at the origin-industry level, by interacting average growth rate in employment, value-added, investment and import share between 2000 and 2003 with a linear time trend. Additional controls include firm age and age squared. Standard errors (in parentheses) are clustered at the origin-industry level.

Table A19: **First Stage with Competition**

VARIABLES	(1) Legal dummies&Market size I Distance	(2) Legal dummies&Market size I Diaspora	(3) Legal dummies&Market size II Distance	(4) Legal dummies&Market size II Diaspora
FLM_{oit}	-0.811	-0.310	-0.530	0.125
	(0.618)	(0.636)	(0.511)	(0.585)
Observations	2,691	2,691	2,691	2,691
R-squared	0.466	0.465	0.466	0.465
Clusters	246	246	246	246
F-stat	1.722	0.238	1.074	0.0454
Mean competition obstacle (percent of firms)	10.24	10.24	10.24	10.24
SD competition obstacle	10.16	10.16	10.16	10.16
Mean Δ competition obstacle	-0.0697	-0.0697	-0.0697	-0.0697
SD Δ competition	10.14	10.14	10.14	10.14
Origin#Industry FE	✓	✓	✓	✓
Year#Origin FE	✓	✓	✓	✓
Year#Industry FE	✓	✓	✓	✓
Linear trends	✓	✓	✓	✓
Level	Industry	Industry	Industry	Industry

Notes: This table presents the first stage of an alternative dependent variable - the share of firms reporting intense product-market competition as an obstacle to doing business. The columns compare first stage results for different variations of the FLM variable. As proximity weights, we use either inverse bilateral distances (Columns 1 and 3) or previous bilateral distribution of emigrants (Columns 2 and 4). Market size represents the *pre-determined* capacity of destination industries to absorb migrants from NMS. Market size I is the number of employees multiplied by the share of firms reporting labor shortages in a given destination industry as of 2000-2003. Market size II is the number of employees multiplied by the share of employed migrants in a given destination industry. All specifications control for origin-industry, origin-year, and industry-year fixed effects. In addition to the fixed effects, we control for differential time trends at the origin-industry level, by interacting average growth rate in employment, value-added, investment and import share between 2000 and 2003 with a linear time trend. Standard errors (in parentheses) are clustered at the origin-industry level.

A.3 Data Appendix

A.3.1 Overview of Migration Data

Data collected from the National Statistical Offices

Austria

Main Association of Austrian Social Security Institutions: posteingang.allgemein@hvb.sozvers.at
Source: Austrian social security data, universe of workers who are subject to social security contributions
Migrants identified by nationality
Available for 2000-2016

Belgium

1)

Belgian statistical office: demos@economie.fgov.be
Source: Statbel (Direction générale Statistique - Statistics Belgium), Labour Force Survey
Migrants identified by country of birth
Available for 2000-2017

2)

Belgian crossroad bank for social security : `https://www.ksz-bcss.fgov.be/en`
Source: administrative data from the Belgian national registry data, universe of workers subject to social security contributions
Migrants identified by country of birth
Available for 2008-2017

Finland

Statistics Finland, Population and Social Statistics: `www.stat.fi`
Source: Universe of workers in Finland based on a compilation from Statistics using different administrative and statistical data
Migrants identified by country of birth
Available for 2000-2016

France

Réseau Quetelet, ADISP `https://quetelet.casd.eu`
Source: Population census `https://www.insee.fr/fr/information/1303686`
Migrants identified by country of birth

Available for 1999, 2006, 2011, 2016 (possible to obtain only for every five years)

Norway

Statistics Norway okonomi@ssb.no
Source: NAV's Employee Register (Aa Register) and A-ordninge, data from coordinated digital collection of employment, income and tax deductions for the Tax Administration, NAV and Statistics Norway
Migrants identified by country of birth
Available for 2000-2018

Sweden

Statistics Sweden, Microdata Unit: `www.scb.se`
Source: administrative registers
Migrants identified by country of birth
Available for 2000-2016

Spain

National Statistics Institute, INE, `https://www.ine.es/en/index.htm`
Source: Labor Force Survey
Migrants identified by nationality
Available for 2006, 2008, 2010, 2012, 2014, 2016

Switzerland

Federal Statistical Office, `https://www.bfs.admin.ch/bfs/en/home.html`
Source: Swiss Labour Force Survey (SAKE)
Migrants identified by country of birth
Available for 2000-2018

United Kingdom

Office for national statistics `https://www.ons.gov.uk`
Source: Annual Population Survey data
Migrants identified by country of birth
Available from 2000 to 2018; the dataset from our request is published here
`https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/`
`\employmentandemployeetypes/adhocs/10617employmentbyeuropeancountryofbirth\`
`ukbetweenmarch2000tofebruary2001andjanuarytodecember2018`

Proxy data for missing migration data

We used Eurostat data on migration stocks at year, destination and origin (country of birth) level (migr_pop3ctb). For missing data, we used another migration dataset at year, destination, and country of citizenship level (migr_pop1ctz). The correlation between stocks of migrants by birth and by nationality is 0.927.

We completed missing Eurostat data from the OECD International Migration Database (Dataset MIG, Stocks of immigrants by country of birth in OECD countries).

For every country-pair, if some internal observations were missing, we linearly interpolated them using the STATA command ipolate.

To "distribute" migrant stocks from each origin country by industry (two-digit level, NACE Rev. 2) in the destination country, we requested Eurostat migration data: by year, destination, region of origin (aggregated to EU3 and EU10), and industry at two-digit level. From these data we obtained the distribution of EU3 and EU10 migrants across industries in old EU member states for each year. The assumption here is that migrants from the same region of origin (EU3 or EU10) work in same industries.

Correlation of the proxy data with non-missing migration data collected from the national statistical offices is 0.72

A.3.2 TFP Calculation Description

Theoretically, TFP is calculated by dividing value-added by the weighted average of labor and capital. When estimating it practically, however, one runs into endogeneity challenges due to the simultaneity of inputs and outputs. The literature on productivity estimations has comprehensively discussed this issue (Olley and Pakes, 1996). When productivity shocks are observed by managers, they strategically choose their inputs, which creates a bias in the estimation due to simultaneity. Olley and Pakes (1996) were the first to introduce a semi-parametric estimation strategy that overcomes the endogeneity by using inputs of capital to proxy for the observed part of the productivity shock. Levinsohn and Petrin (2003) further develop the method and make it more feasible to estimate it empirically by using variable inputs such as materials as a proxy for the observed part of the productivity shock. As we observe materials in our dataset, we can apply the Levinson and Petrin methodology. We apply the *prodest* command in STATA to easily implement it (Mollisi and Rovigatti, 2017). This methodology has been extensively used in the literature (Blalock and Gertler, 2004; Topalova and Khandelwal, 2011) and further developed by Wooldridge (2009). We check the robustness of our estimations using the latter methodology as well.

A.3.3 Orbis Dataset Description

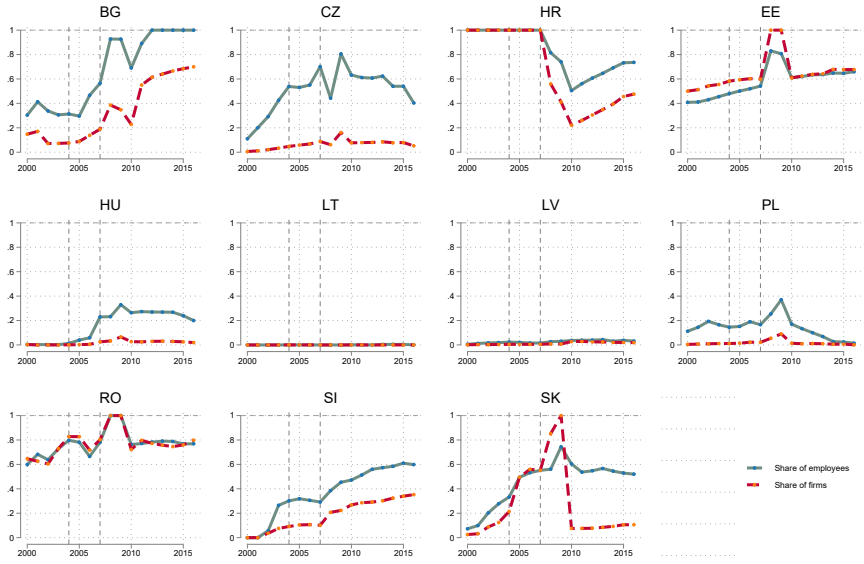
Orbis is a commercial firm-level database maintained by Bureau van Dijck (BvDEP). It is a collection of business statements (in particular, balance sheets and profit and loss statements), ratings and news of firms (mostly) from the private sector. Orbis contains financial information for more than 375 million companies. Although the Orbis database claims to cover all countries, the coverage varies strongly not only from country to country, but also by sector, time and type of firm information available.

The Orbis dataset is not representative for the entire population of firms. Bajgar et al. (2020) find that the average Orbis firm is larger, older and more productive. However, restricting the sample to only the best-covered countries and imputing missing values improves the representativeness. To check representativeness of our dataset, we aggregate the number of employees and the number of firms per origin country and year and compare these indicators with the aggregates from the Eurostat Structural Business Statistics (Figure A7). Coverage varies by country: while it is rather good for Bulgaria, Estonia and Romania, we have only few firms in Latvia, Lithuania, and Hungary. It can be also seen that on average firms are larger: with a smaller share of firms we cover larger shares of employees. Coverage also varies over time. Therefore, we use Orbis sample to study within-firm effects of emigration.

The data have been successfully used in academic economic research, mainly by papers studying multinational enterprises (Egger et al. (2009); Beer and Loeprick (2015)). Moreover, Gal (2013) uses Orbis data to measure total factor productivity at the firm level. More information on the methodology and access formalities can be found online.[42]

[42] https://www.bvdinfo.com/en-us/our-products/data/international/orbissecondaryMenuAnchor0/

Figure A7: **Comparison of Orbis and Eurostat Aggregate Figures**



Notes: This figure compares aggregate number of firms and number of employees per NMS and year between our Orbis sample and Eurostat Structural Business Statistics. A dashed line shows the share of firms from Eurostat that we cover with Orbis data and a continuous line shows the share of covered employees. Eurostat data for Croatia (HR) is available starting from 2008.

A.3.4 Correspondence Tables

One challenge with the independent variables at the two-digit industry level (migration data, training, structural business statistics, etc.) arises from the change in NACE classification (Revision 2 replaced Revision 1 in 2008). Some of our data are thus available only in Revision 1 and some only in Revision 2. At a two-digit level, we run into a problem of the many-to-many relation (several NACE Revision 1 codes can potentially map into several NACE Revision 2 codes).[43] We overcame this problem by creating a country-specific conversion matrix using Orbis data for 2009, where we can observe both Revision 1 and Revision 2 NACE codes for the same firm. For each NACE Revision 1 two-digit code, we obtain a corresponding weight (share) for each NACE Revision 2 code. The weights range between zero and one, sum to one, and are equal to the share of firms reporting a given Revision 2 code in the total number of firms with a given Revision 1 code.

[43] Correspondence tables are well-defined at a four-digit level, but we do not always have data available at that level.

A.4 Anecdotes from the Media