CHAPTER 4

ONLINE ANNEX

Annex 4.1: List of economies

Table 1: List of Economies and Regions

East Asia & Pacific

American Samoa	Japan	Myanmar	Solomon Islands
Australia	Kiribati	Nauru	Taiwan Province of China
Brunei Darussalam	Korea, Dem. People's Rep.	New Caledonia	Thailand
Cambodia	Korea	New Zealand	Timor-Leste, Dem. Rep. of
China	Lao P.D.R.	Northern Mariana Islands	Tonga
Fiji	Macao SAR	Palau	Tuvalu
French Polynesia	Malaysia	Papua New Guinea	Vanuatu
Guam	Marshall Islands	Philippines	Vietnam
Hong Kong SAR	Micronesia, Fed. States of	Samoa	
Indonesia	Mongolia	Singapore	

Europe & Central Asia

Albania	Faroe Islands	Kyrgyz Republic	San Marino
Andorra	Finland	Latvia	Serbia
Armenia	France	Liechtenstein	Slovak Republic
Austria	Georgia	Lithuania	Slovenia
Azerbaijan	Germany	Luxembourg	Spain
Belarus	Gibraltar	Moldova	Sweden
Belgium	Greece	Monaco	Switzerland
Bosnia and Herzegovina	Greenland	Montenegro, Rep. of	Tajikistan
Bulgaria	Hungary	Netherlands	Turkey
Channel Islands	Iceland	FYR Macedonia	Turkmenistan
Croatia	Ireland	Norway	Ukraine
Cyprus	Isle of Man	Poland	United Kingdom
Czech Republic	Italy	Portugal	Uzbekistan
Denmark	Kazakhstan	Romania	
Estonia	Kosovo	Russia	

Latin America & Caribbean

Antigua and Barbuda	Colombia	Haiti	St. Lucia
Argentina	Costa Rica	Honduras	St. Martin (French part)
Aruba	Cuba	Jamaica	St. Vincent and the Grenadines
Bahamas, The	Curaçao	Mexico	Suriname
Barbados	Dominica	Nicaragua	Trinidad and Tobago
Belize	Dominican Republic	Panama	Turks and Caicos Islands
Bolivia	Ecuador	Paraguay	Uruguay
Brazil	El Salvador	Peru	Venezuela
British Virgin Islands	Grenada	Puerto Rico	Virgin Islands (U.S.)
Cayman Islands	Guatemala	Sint Maarten (Dutch part)	
Chile	Guyana	St. Kitts and Nevis	

Middle East & North Africa

Algeria	Israel	Morocco	United Arab Emirates
Bahrain	Jordan	Oman	West Bank and Gaza
Djibouti	Kuwait	Qatar	Yemen
Egypt	Lebanon	Saudi Arabia	
Iran	Libya	Syria	
Iraq	Malta	Tunisia	

North America

Bermuda
Canada
United States

South Asia

Afghanistan	India	Pakistan
Bangladesh	Maldives	Sri Lanka
Bhutan	Nepal	

Sub-Saharan Africa

Angola	Côte d'Ivoire	Liberia	Senegal
Benin	Equatorial Guinea	Madagascar	Seychelles
Botswana	Eritrea	Malawi	Sierra Leone
Burkina Faso	Eswatini	Mali	Somalia
Burundi	Ethiopia	Mauritania	South Africa
Cabo Verde	Gabon	Mauritius	South Sudan
Cameroon	Gambia, The	Mozambique	Sudan
Central African Republic	Ghana	Namibia	Tanzania
Chad	Guinea	Niger	Togo
Comoros	Guinea-Bissau	Nigeria	Uganda
Democratic Republic of the Congo	Kenya	Rwanda	Zambia
Congo, Republic of	Lesotho	São Tomé and Príncipe	Zimbabwe

Source: World Bank.

Annex 4.2: The Drivers of Migration

A. Model Specification and Estimation for the Section on the Drivers of Migration

1. This chapter uses a standard gravity-type model to analyze the drivers of the numbers of immigrants and emigrants jointly. The micro-foundations of the gravity model for migration are based on a representative individual who maximizes her utility by choosing between staying home or migrating to one of many destination countries. The individual weighs up the benefits of migrating, especially the wage differential between destination and origin countries, and the various costs of migrating. The aggregate choices of many individuals in an origin country result in migration rates per person that follow a multinomial distribution across destination countries. For a detailed explanation, see Beine et al (2016).

2. Using a standard transformation (Baker, 1994), the number of migrants can then be modelled using a (conditional) Poisson distribution. The mean of the Poisson distribution depends on drivers, which are the benefits and costs of migrating that appear in the representative individual's utility function above. The expected number of migrants, conditional on the drivers, is then

$$E\left(M_{j,k,t} \middle| x_{j,k,1}, \dots, x_{j,k,T}, \eta_k^d, \theta_t\right) = \exp\left(x'_{j,k,t}\beta + \eta_k^d + \theta_t\right) \qquad (1)$$

where $M_{j,k,t}$ is the gross flow of migrants from origin country j to destination country k in period t and η_k^d, θ_t are the destination and time fixed effects respectively. The column vector $x_{j,k,t}$ contains the drivers of interest. For the baseline specification, these include:

a. the logs of average GDP per person in the preceding period, in the destination country k and origin country j, in thousands of 2011 PPP international dollars. The square of the log of origin country GDP per person is also included to capture poverty effects;

b. the log of average of population of the origin country j over in the preceding period;

c. the intensity of war measured on a scale from zero to ten (explained below), including ethnic, civil and international war episodes;

d. the log of the distance in kilometers between origin and destination countries; and

e. indicator variables for whether countries k and j share a former colonial relationship, a common border, and a common language.

3. The extended specification tests for further potential drivers, including:

a. contemporaneous average annual growth of real GDP per person in the period

b. contemporaneous 5-year ahead expectations of future economic growth at destination and origin

c. contemporaneous unemployment rates at origin and destination;

d. the log of the number of migrants from the origin country j in the destination country k at the end of the previous period;

e. measures of the tightness of entry, control and integration immigration policies in the destination country in the contemporaneous period;

f. average over the contemporaneous period of temperature in degrees Celsius at the origin country;

g. numbers of natural disasters in the origin country over the contemporaneous period, including[1] droughts, earthquakes, epidemics, floods, insect infestations, landslides, storms, extreme temperature events, volcanic eruptions and miscellaneous events;

h. indicators of whether a debt, banking or currency crisis occurred in the origin country at any time in the contemporaneous period.

4. Model (1) is estimated by Poisson quasi-maximum likelihood, using the fixed effects Poisson estimator of Hausman et al. (1984). The key assumption for consistency is strict exogeneity of the drivers $x_{j,k,t}$ (see Cameron and Trivedi, 2005, section 23.7).

B. Data

5. Data for international migrant stocks are taken from the 2019 edition of those published by the UN Department of Economic and Social Affairs (DESA). These data show the numbers of international migrants for every origin and destination country pair in the world, every 5 years from 1990 to 2015. Data are also available for 2019, but these data are not used in the estimation of the model of drivers of migration because the four-year gap between 2015 and 2019 is inconsistent with the five-year frequency of the analysis in this section. The data primarily come from the national censuses of each destination country (or area), but they are complemented by population registers, nationally representative surveys, and numbers of refugees and asylum seekers from UN agencies.

6. The data define an international migrant as a foreign-born person, but for about one-fifth of all countries, an international migrant is defined as a foreign citizen. The origin of the migrant is then their place of birth or country of citizenship respectively. This lack of consistency in definitions harms the quality of the data and is a caveat to this analysis.

7. The data on international migrant stocks have been estimated by the compilers in at least three important ways: (i) refugees and asylum seekers are added to numbers of migrants where they would otherwise be excluded; (ii) migrants are distributed among specific countries when the data provide information only for broadly-defined geographic regions; and (iii) numbers of migrants are also interpolated in years where no data sources are available, which frequently occurs for example between census years.

8. The analysis here of numbers of migrants sticks to the micro-foundations of model (1) by using migration flows as the dependent variable. It follows two methods used in the literature to

[1] The types of natural disasters follow those used in Beine and Parsons (2015).

derive migration flows from migrant stocks. The first, called *left-censored migration flows*, uses changes in stocks and sets all negative changes to zero. The second, called *return-adjusted flows*, treats all negative changes as return migration, by adding them to migration flows in the opposite direction. The reason for eliminating negative net migration flows is to relate the estimations to the theoretical model, which is defined for gross migration flows. Therefore, the interpretation of the left-censored and return-adjusted flows lies somewhere between net flows and gross flows. The baseline estimations use 5-year flows, but many countries only conduct censuses every 10 years, so this appendix discusses how the results change when using 10-year flows instead.

9. Data on the education composition of migrants come from the DIOC-E database (release 1.0), covering OECD destination countries and most origin countries. Only the data for the 2010 census round are used, and the relevant population is all ages.

10. Data on the intensity of war are obtained from the Major Episodes of Political Violence database produced by the Center for Systemic Peace. The data cover conflict episodes with a systematic and sustained use of lethal violence by organized groups that resulted in at least 500 directly-related deaths over the episode, and they measure intensity on a scale from zero to ten, where zero is no conflict and ten is the most intense. A conflict of intensity 1 denotes sporadic or expressive political violence (e.g. U.S. 1965, Argentina 1982) and intensity 10 denotes extermination and annihilation (e.g. WWII).

11. Demographic data are obtained from the UN DESA. They include population and numbers of people by age group. From this, the chapter constructs numbers of young people as the number of people between the ages of 15 and 29, following Clark, Hatton and Williamson (2002) and Hatton and Williamson (2003). Data on banking, currency and debt crises are obtained from the database of Laeven and Valencia (2013). Data on migration policies are obtained from both the Determinants of International Migration (DEMIG) project and the Immigration Policies in Comparison (IMPIC) project (Helbling et al., 2017). The DEMIG data measure the numbers of policy tightenings or loosenings during each year, while the IMPIC data measure the tightness of policy at each point in time on a scale from zero to one. Data on natural disasters are obtained from the International Disasters Database published by the Centre for Research on the Epidemiology of Disasters. Data on temperatures come from annual average country temperatures in the CRU CY dataset (version 4.03) produced by the University of East Anglia's Climate Research Unit.

Table 1. Baseline Estimations for 5-Year Left-Censored Migration Flows

	world	EM→EM	EM→AE	AE→AE
lag income destination	1.07***	0.78***	1.12*	0.44
	(0.27)	*(0.3)*	*(0.51)*	*(0.55)*
lag income origin	0.76***	0.54***	0.89***	1.5
	(0.11)	*(0.15)*	*(0.18)*	*(1.69)*
lag $(\text{income})^2$ origin	-0.29***	-0.33***	-0.23***	-0.50.
	(0.03)	*(0.05)*	*(0.05)*	*(0.29)*
lag (inc. gap) x (young pop. origin)	0.02	0.10***	0.01	0.07
	(0.02)	*(0.03)*	*(0.03)*	*(0.08)*
lag population origin	0.52***	0.22***	0.61***	0.61***
	(0.04)	*(0.06)*	*(0.07)*	*(0.04)*
lag young share origin	0.634	1.956	-0.383	-0.283
	(0.57)	*(1.24)*	*(0.42)*	*(0.61)*
war origin	0.12**	0.23***	-0.004	-0.14
	(0.05)	*(0.07)*	*(0.02)*	*(0.10)*
distance	-1.23***	-1.18***	-1.14***	-0.80***
	(0.08)	*(0.17)*	*(0.09)*	*(0.08)*
common border	1.17***	1.28***	1.17***	0.14
	(0.16)	*(0.15)*	*(0.21)*	*(0.21)*
common language	0.31***	0.11	0.82***	1.24***
	(0.12)	*(0.18)*	*(0.12)*	*(0.16)*
colonial link	1.12***	2.35***	1.07***	-0.15
	(0.12)	*(0.26)*	*(0.14)*	*(0.29)*
turning point	3,747	2,288	6,930	4,459
# observations	42,592	15,460	14,540	4,753
# destinations	173	135	36	36
# origins	153	121	121	32
R^2 (percent)	42	35	73	26

Notes:

All specifications include destination country fixed effects and period fixed effects, but not origin country fixed effects.

Standard errors for coefficients clustered by origin--period are in parentheses.

Significance stars and p-values: *** < 0.01 < ** < 0.025 < * < 0.05 < . < 0.1

The turning point is the origin country income (in 2011 PPP US$ per person) at which the marginal effect of origin country income is zero, assuming a zero income gap between origin and destination.

The following 'corridors' are used: EM->AE denotes from EMDEs to AEs, AE->AE denotes from AEs to AEs , EM->EM denotes from EMDEs to EMDEs, and AE->EM denotes from AEs to EMDEs.

C. Estimation Results

12. Table 1 shows the baseline estimation results, which are discussed in the chapter's main text and only briefly elaborated on here. It is useful to combine the interpretation of Table 1 with references to Table 2, which shows what happens to the model findings when changes are made to the dependent variable, sample period and standard errors.

13. Consider first the effects of war. A one unit increase in war intensity on this scale corresponds, for example, to going from a conflict of 3-10 thousand deaths to one of 10-50 thousand deaths (intensity 2 to intensity 3). The model on the world sample suggests that emigration flows increase by 12 percent with this increase in war intensity, while between EMDEs, emigration flows increase by 25 percent. The effects of war seem relatively short-lived, with many migrants returning between 5 and 10 years after the intensification of war, as evidenced by the lack of statistical significance in models of 10-year flows and of migrant stocks (Table 2).

14. As expected, there is less migration between countries that are further apart. Interestingly however, migration flows from EMDE countries are inversely proportional to distance, which means that migration flows from an EMDE country to any other country are half as large when those two countries are twice as far apart. The estimates also suggest that migration flows are about 3 times as large between countries that share a border or colonial relationship. They are about 1.5 times as large between countries that share a common language.

15. Various forms for temperature were tested as potential drivers of international migration,[2] including deviations from ideal temperatures of 25 Celsius, non-linearities, and interactions with agricultural production. Consistent with Beine and Parsons (2015) and Cattaneo and Peri (2016), temperature mostly did not show robust effects, holding incomes constant. However, there was one exception. Increases in average temperature were associated with less international emigration from poor countries, holding their income constant. This "climate poverty trap" channel operates in addition to the one stemming from the reduction in income due to rising temperatures in already poor and hot countries (October 2017 *World Economic Outlook*). Interestingly, this channel holds for migration between EMDEs but not from EMDEs to AEs. It is also long-lasting.

16. Controlling for its effect on origin country income, a currency crisis leads to about 29 percent more emigration flow over a 5-year period. Effects over decades (Table 2) are even stronger, suggesting that emigration increases between 5 and 10 years after a currency crisis. However, the effects of crises are not permanent, and migrants eventually return, as suggested by the lack of effect on migrant stocks (Table 2).

17. Table 3 shows the estimation results for the drivers of the composition of migration. These are interpreted in the chapter's main text so only the details are explained here. The dependent variable is the skill composition of the migrant stock in 2010, which is defined as

[2] As noted in the main text, temperature has been found to be a driver of internal migration (Rigaud et al., 2018).

$$\log\frac{N_{j,k}^h}{N_{j,k}^l} - \log\frac{N_j^h}{N_j^l}$$

where $N_{j,k}^h$ is the number of high-skill migrants from country j in country k, and $N_{j,k}^l$ is the corresponding number of low-skilled. N_j^h and N_j^l are the numbers of high-skill and low-skill workers in the population. When the high-skilled are taken as those with a tertiary education and the low-skilled are taken as those with a primary education, this variable is the measure of skill selection in Grogger and Hanson (2011). This dependent variable is denoted by "3/1" in Table 3. When the high- and low-skilled are taken as the tertiary- and secondary-educated respectively, the dependent variable is denoted as "3/2". Similarly, when the skill composition compares the secondary- to primary-educated, it is denoted "2/1". The skill selection equations are estimated by ordinary least squares.

Table 2. Sensitivity of Parameter Estimates in the Statistical Models of the Numbers of Migrants

Finding	Dependent variable				Sample		White (1980) standard errors
	5-year RA flows /1	10-year LC flows /1	10-year RA flows /1	stocks (5-year frequency)	early (1990-2000)	late (2000-15)	
Origin income turning point around $4,000 p.p.	Yes	Yes	Yes	Closer to $5,000 p.p.	Closer to $3,000 p.p.	Yes	Yes
Positive effect of destination income	Yes	Yes /2	Yes, but only at 10% level /2	No	Yes	Yes, but only after controlling for crises	Yes
Positive effect of origin population	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Positive interaction between income and young population	Yes, but only at 10% level	Yes /3	Yes /3	Yes	Yes, but only at 10% level	Yes	Yes
Distance, common border, common language, colonial link	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Positive effect of war	Yes	Yes, but not statistically significant	Yes, but not statistically significant	Yes, but not statistically significant	Yes	Yes, but not statistically significant	Yes
Positive network effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Positive effect of extreme temperatures	Yes	Yes	Yes	Yes	Yes	Yes, but not statistically significant	Yes
Poverty trap in temperature	Yes	Yes	Yes	Yes	Yes	Yes, but only at 10% level	Yes
Positive effect of currency crises	Yes, but only at 10% level	Yes	Yes	No, only joint with a bank crisis	Yes, but only at 10% level	No, only debt crisis	Yes

Notes:

1/ LC denotes left-censored flows and RA denotes return-adjusted flows, as explained in the text.

2/ Using contemporaneous income, rather than lagged income.

3/ Only when controlling for lagged income at orign and destination.

18. A key potential driver of the composition of migrants could be the skill premium. In the absence of skill premium data on a broad sample of countries, this chapter follows Grogger and Hanson (2011) in approximating the skill premium between tertiary- and primary-educated workers by the difference between the 80th and 20th percentiles of the distribution of real (2011 PPP) GDP per person in each country. In turn, these percentiles are estimated from income inequality data under the assumption that real GDP per person is lognormally distributed in the population. It uses the median of the distribution of real GDP per person as the earnings of the secondary-educated, from which skill premia between the tertiary- and secondary-educated and between the secondary- and primary-educated can be computed. These skill premia are computed both as dollar differences and as percentage differences, but the latter is dropped due to difficulties of interpreting the results. One drawback of the approach to estimating the skill premium is that it induces correlation between the skill premium and destination income, which makes the effect of destination income difficult to identify, so it is excluded from the models below.

19. As explained in the main text, the education-specific skill premium is estimated to be a positive and statistically significant driver of the education of migrants (Columns (1)-(3) of Table 3). Columns (4) and (5) add countries' exposure to automation as a potential driver of the skill composition of migrants. The exposure to automation is measured though the country's routine task intensity (RTI) index, which is explained in the main text (Box 4.1) and aggregated up from sector-level indices. The index is only available for a small set of economies, so the sample size drops dramatically, and hence the results are merely suggestive. Column (4) of Table 3 shows a negative coefficient on the RTI, suggesting that destination countries with more routine jobs attract relatively more secondary- than tertiary-educated people, and origin countries with fewer routine jobs send relatively more secondary- than tertiary-educated people abroad. However, the coefficient is not statistically significant, so this tendency may not be strong. Similarly, Column (5) shows that destination countries with more routine jobs tend to attract more secondary- than primary-educated people (and vice versa for origin countries' emigrants), and this effect is statistically significant. Together, the results of Columns (4) and (5) are consistent with stronger demand for secondary-educated workers in countries with more routine jobs.

D. Decomposition of Changes over Time

20. The estimated gravity model can be used to decompose the changes over time in migrant stocks into the contributions from each of the explanatory variables (drivers) in the model. Figure 11 in the main text shows this decomposition.

21. The decomposition starts from the model for migrant stocks $\mathrm{N}_{\mathrm{j,k,t}}$, for which the conditional mean specification is

$$\hat{N}_{j,k,t} \coloneqq E(N_{j,k,t}|x_{j,k,1},\dots,x_{j,k,T},\eta_k^d,\theta_t) = \exp\left(x_{j,k,t}^{'}\beta + \eta_k^d + \theta_t\right). \qquad (2)$$

Let the fitted values from the estimated model be denoted $\widehat{N}_{j,k,t}$, and write the identity $N_{j,k,t} = \widehat{N}_{j,k,t}e_{j,k,t}$, where $e_{j,k,t}$ is the (multiplicative) model residual. From this, the change over time in the stock of migrants from EMDEs in AEs is

Table 3. Skill Selection Regressions: OLS on Cross-Section in 2010

	(1)	(2)	(3)	(4)	(5)
	3/1	3/2	2/1	3/2	2/1
tertiary/primary skill premium (dest.-orig.)	0.018***				
	(0.002)				
tertiary/secondary skill premium (dest.-orig.)		0.010***		0.032***	
		(0.002)		*(0.005)*	
secondary/primary skill premium (dest.-orig.)			0.028***		0.039***
			(0.005)		*(0.01)*
routine task intensity (dest.-orig.)				-0.084	1.369***
				(0.44)	*(0.522)*
lag migrants	-0.103***	-0.077***	-0.021.	-0.087***	-0.026
	(0.015)	*(0.008)*	*(0.012)*	*(0.021)*	*(0.032)*
distance	0.259***	0.161***	0.113***	-0.110*	0.079
	(0.036)	*(0.024)*	*(0.029)*	*(0.056)*	*(0.048)*
common border	-0.737***	-0.313***	-0.443***	-0.484***	-0.284*
	(0.136)	*(0.079)*	*(0.115)*	*(0.103)*	*(0.137)*
common language	0.314***	0.292***	0.066	0.265***	0.323***
	(0.078)	*(0.053)*	*(0.061)*	*(0.094)*	*(0.106)*
colonial link	0.185	0.356***	-0.192		
	(0.219)	*(0.091)*	*(0.158)*		
# observations	3,135	3,264	3,151	428	424
# destinations	75	75	75	23	23
# origins	123	124	124	23	23
R^2 (percent)	60	47	56	40	39

Notes:

The dependent variable is indicated in the second row. "3/1" denotes the skill selection measure between those with tertiary and primary education, as explained in the text, from Grogger and Hanson (2011). The other dependent variables, "3/2" and "2/1", are measures of skill selection based on those with secondary education, as explained in the text.

Standard errors for coefficients clustered by origin--period are in parentheses.

Significance stars and p-values: *** < 0.01 < ** < 0.025 < * < 0.05 < . < 0.1

$$
\begin{aligned}
N_{AE,2015} - N_{EMDE,1990} \quad &= \quad \sum_{k \in AE} \sum_{j \in EMDE} \sum_{\tau=1}^{5} \left(N_{j,k,1990+5\tau} - N_{j,k,1990+5(\tau-1)} \right) \\
&\approx \sum_{k \in AE} \sum_{j \in EMDE} \sum_{\tau=1}^{5} N_{j,k,1990+5(\tau-1)} \log\left(N_{j,k,1990+5\tau} / N_{j,k,1990+5(\tau-1)} \right), \\
&\approx \sum_{k \in AE} \sum_{j \in EMDE} \sum_{\tau=1}^{5} \widehat{N}_{j,k,1990+5(\tau-1)} \log\left(\widehat{N}_{j,k,1990+5\tau} / \widehat{N}_{j,k,1990+5(\tau-1)} \right)
\end{aligned}
$$

which relies on the log-linear approximation $\log(1 + x) \approx x$ and on the quality of the fit of the model, $N_{j,k,t} \approx \widehat{N}_{j,k,t}$. Each of these is a source of error in the decomposition that is quantified below. Then use the model specification (2) to replace $\log\left(\widehat{N}_{j,k,1990+5\tau} / \widehat{N}_{j,k,1990+5(\tau-1)}\right)$ by

$$
\left(x_{j,k,1990+5\tau} - x_{j,k,1990+5(\tau-1)} \right)' \beta + \theta_{1990+5\tau} - \theta_{1990+5(\tau-1)}.
$$

If the vector $x_{j,k,t}$ contains four time-varying explanatory variables[3]

$$
x'_{j,k,t} = \begin{pmatrix} y^d_{j,k,t} & y^o_{j,k,t} & \left(y^o_{j,k,t}\right)^2 & p_{j,k,t} \end{pmatrix},
$$

then one can reverse the order of summation above to find that the change in the stock of migrants $N_{AE,2015} - N_{EMDE,1990}$ can be written as the sum of the following four terms:[4]

Contribution of destination income	$\sum_{k \in AE} \sum_{j \in EMDE} \sum_{\tau=1}^{5} \widehat{N}_{j,k,1990+5(\tau-1)} \left(y^d_{j,k,1990+5\tau} - y^d_{j,k,1990+5(\tau-1)} \right) \beta_1$
Contribution of origin income (in levels and squares)	$\sum_{k \in AE} \sum_{j \in EMDE} \sum_{\tau=1}^{5} \widehat{N}_{j,k,1990+5(\tau-1)} \left[\left(y^o_{j,k,1990+5\tau} - y^o_{j,k,1990+5(\tau-1)} \right) \beta_2 + \left(\left(y^o_{j,k,1990+5\tau} \right)^2 - \left(y^o_{j,k,1990+5(\tau-1)} \right)^2 \right) \beta_3 \right]$

[3] Because they are differenced across time, the time-invariant explanatory variables like distance, common border, common language and colonial link are eliminated. Income at destination, income at origin and origin population are denoted $y^d_{j,k,t}$, $y^o_{j,k,t}$ and $p_{j,k,t}$ respectively.

[4] This decomposition relies on the notation $\beta' = (\beta_1 \quad \beta_2 \quad \beta_3 \quad \beta_4)$.

Contribution of origin population	$\sum_{k\in AE}\sum_{j\in EMDE}\sum_{\tau=1}^{5}\widehat{N}_{j,k,1990+5(\tau-1)}\big(p_{j,k,1990+5\tau}-p_{j,k,1990+5(\tau-1)}\big)\beta_4$
Contribution of time fixed effect	$\theta_{1990+5\tau}-\theta_{1990+5(\tau-1)}$

22. This is the approach used to produce Figure 4.11, where the formulae for AE—AE and EMDE—EMDE migration are analogous to the formula above for EMDE—AE migration. However, two modifications are made to this method. First, coefficients are allowed to be corridor-specific, which makes this historical composition comparable with future migration scenarios (where the coefficients are also allowed to vary by corridor). Corridor-specific coefficients have the side-effect of improving the fit of the model in-sample. Second, coefficients that are not statistically significant at the 5 percent level are set to zero, to avoid statistically insignificant coefficients with the incorrect sign contaminating in the interpretation of the decomposition. Setting coefficients to zero introduces another source of error in the model decomposition that is termed "variable selection error".

23. Table 4 shows the aggregate size of the errors in decomposition from each source along each migration corridor. It also shows the time trend, which could be thought of either as an unexplained component of the model or as a proxy for many known global factors driving migration, like average travel costs. The largest sources of error in the decomposition are due to variable selection and log-linear approximation. The errors due to missing data and model fit are smaller.

Table 4. Errors and unexplained variation in the time decomposition of migration stocks, 1990--2015. (In millions of new migrants)

	source of error/unexplained variation				
corridor	time trend	missing data	estimation	variable selection	approx-imation
EM→EM	27.0	0.2	-3.7	-2.6	0.9
EM→AE	7.9	2.6	0.6	-0.1	5.7
AE→AE	4.0	0.6	1.2	5.6	0.8

E. Model Specification and Estimation of Migration Scenarios

The specification and estimation method of the gravity model used in the section of future migration scenarios are essentially the same as the ones presented above, with two differences. First, the model is now estimated using bilateral stocks (rather than flows). Only stocks up to 2010 are used given the relatively more reliable information that is contained in the decadal census data. Second a full set of corridor-specific dummies are interacted with all the drivers. Results are broadly in line with the corresponding corridor estimates discussed above. As mentioned before, war/conflict drivers are not significant in stock regressions. Finally, income converge for the baseline scenario assume that income per capita grows at the same rate as in the

United States (zero convergence) if the converge rate form the October 2019 *World Economic Outlook* is negative.

Annex 4.3: The Impact of Large Immigration Waves

24. This section describes the estimation procedure and robustness exercises for section on the impact of large immigration waves examined using local projections methods (Jorda, 2005). This approach captures the dynamic effect over an *h* year horizon of the shocks. As discussed in the chapter's main text, prior to estimating the local projections model, instrumental variables are constructed to ensure exogeneity of migration flows to the various macroeconomic variables of interest and shock episodes are defined.

A. Data

25. Annual data for migration flows is from the OECD, UN, and World Bank. Annual refugee stock data is taken from the UNHCR. The refugee variable is defined as the sum of individuals classified by the UNHCR as refugees, asylum seekers, and other persons of concern. Macroeconomic variables are taken from various sources: output, consumption, capital, and total productivity are from the Penn World Tables, aggregate employment variables are from the IMF's WEO database, and native employment variables from the OECD. The real sector and financial sectors reform indices were provided by Alesina, Furceri, Ostry, Papageorgiou and Quinn (2020). The immigration integration restrictiveness variable is from the Immigration Policies in Comparison (IMPIC) project (reclassified according to Schmid and Helbling 2016). These variables take on the value from zero to one, with one representing more restrictiveness. Spending per capita on active labor market policies is from the OECD and spending on adult education and vocational training is measured as a share of GDP and taken from the IMF's Government Financial Statistics database (classified as "education not definable by level" and defined as education not definable by level includes education programs, generally for adults, that do not require any special prior instruction, particularly vocational training and cultural development). Finally, additional data on bilateral distance between countries and contiguity, used to construct the refugee instrumental variable, is from the CEPII.

B. Instrumental Variable

26. In order to avoid reverse causality between the macroeconomic dependent variables and migration and refugee shocks, exogenous instrumental variables (IVs) are created for both shocks. The IV for migration flows is constructed on the basis of a shift-share instrument following Card (2009) (among many others):

$$\widetilde{m}_{it} = \sum_{o} \frac{M_{oit^o}}{M_{ot^o}} \Delta M_{ot}$$

Where M_{oit^o}/M_{ot^o} is the share of the stock of migrants from origin *o* in destination *i* at historical time t^o, and $t^o = t-5{:}\, t-10$. The range of time t^o is due to migration stock data being available only at five-year intervals. This is multiplied by the total outflow of migrants, ΔM_{ot}, from origin *o* in year *t*.

In the case of refugees, the proxy is defined differently because, as discussed in the chapter's main text, refugees are often forced to flee suddenly and are not necessarily given a choice of their destination country. Instead, refugee flows are more likely determined on the basis of proximity. The proxy is thus defined based on a combination of the historical flows between origin and host countries and proximity variables from the gravity in migration literature, as:

$$\tilde{r}_{it} = \sum_{o} \frac{M_{iot}^{\alpha} D_{iot}^{\gamma} e^{\zeta \cdot contig_{oi}}}{\sum_i M_{iot}^{\alpha} D_{iot}^{\gamma} e^{\zeta \cdot contig_{oi}}}$$

Where the coefficients $\alpha = 0.4$, $\gamma = -0.8$, and $\zeta = 0.6$ are drawn from the gravity migration literature (see Beine, Bertoli, Fernandez-Huertas Moraga (2016) for a review of the literature). The instruments have a relatively high correlation with the actual flows, for migrants is 83 percent and for refugees is 52 percent over the shock episodes.

C. Shock Selection

27. As discussed in the chapter's main text, shock episodes are identified for both migration and refugee inflows based on historical and future trends. For migration, an episode is considered a shock if the annual inflow (as a share of population) is greater than the country's median inflow during the period 1980-2018 and is also greater than the median inflow (relative to the recipient country's population) experienced by OECD countries during the previous five-year period and the following five-year period. Refugee shocks are defined as an inflow (as a share of population) that is within the country's top 10th percentile of inflows during the period 1980-2018 and is also greater than the top 10th percentile (relative to the recipient country's population) experienced by all countries in the world during the five-year period and the following five-year period. Finally, to avoid including episodes characterized by sudden reversals, the refugee inflow shock must be sustained for at least two consecutive years. Inflow values are set to zero if they do not meet the shock criterion.

D. Estimation

28. With instrumental variables in hand and shock episodes identified, the estimating equation to determine the impact of a migration shock on macroeconomic outcomes follows the local projections methodology (Jorda, 2005):

$$y_{i,t+h} - y_{i,t-1} = \alpha_i^h + \gamma_t^h + \beta_1^h \frac{\Delta Mig_{it}}{E_{i,t-1}} + \varepsilon_{it}^h \quad (1)$$

29. Where $y_{i,t}$ are the macroeconomic outcome variables of interest: output, total employment, native employment, total labor force, native labor force, labor productivity, total factor productivity (TFP) (all in logs), capital output ratio, unemployment rate, and native unemployment rate (all as level of ratio). The independent variable is the shock, measured as migration inflows ($\Delta Mig_{i,t}$) relative to the previous period's employment level ($E_{i,t-1}$). This is the relevant ratio to capture the impact of migrants, given the majority are of working age and able to enter the labor force relatively rapidly. For the study of refugees, the independent variable of interest is replaced with $\Delta Ref_{i,t}/E_{i,t-1}$, where the numerator is the annual change in refugee stock (derived from a difference in stocks, due to data availability). Country (α_i^h) and time (γ_t^h) fixed effects are also included to capture time-invariant country-specific factors and global shocks that could affect macro outcomes.

30. The model is estimated via two stage least squares where the change in migration inflows is instrumented with the IV for migration relative to employment ($\Delta\widetilde{m}_{it}/E_{i,t-1}$), and in the case of refugees the inflow is instrumented with the change in the IV to employment ($\Delta\tilde{r}_{it}/E_{i,t-1}$). In all cases the IV is winsorized at the top one percent (upper bound only) to account for extreme values created by our proxy due to data limitation, abnormal one-off flows, or other factors that overly distort the IV. F-statistics suggest are typically above 10, suggesting the instruments have sufficient power.

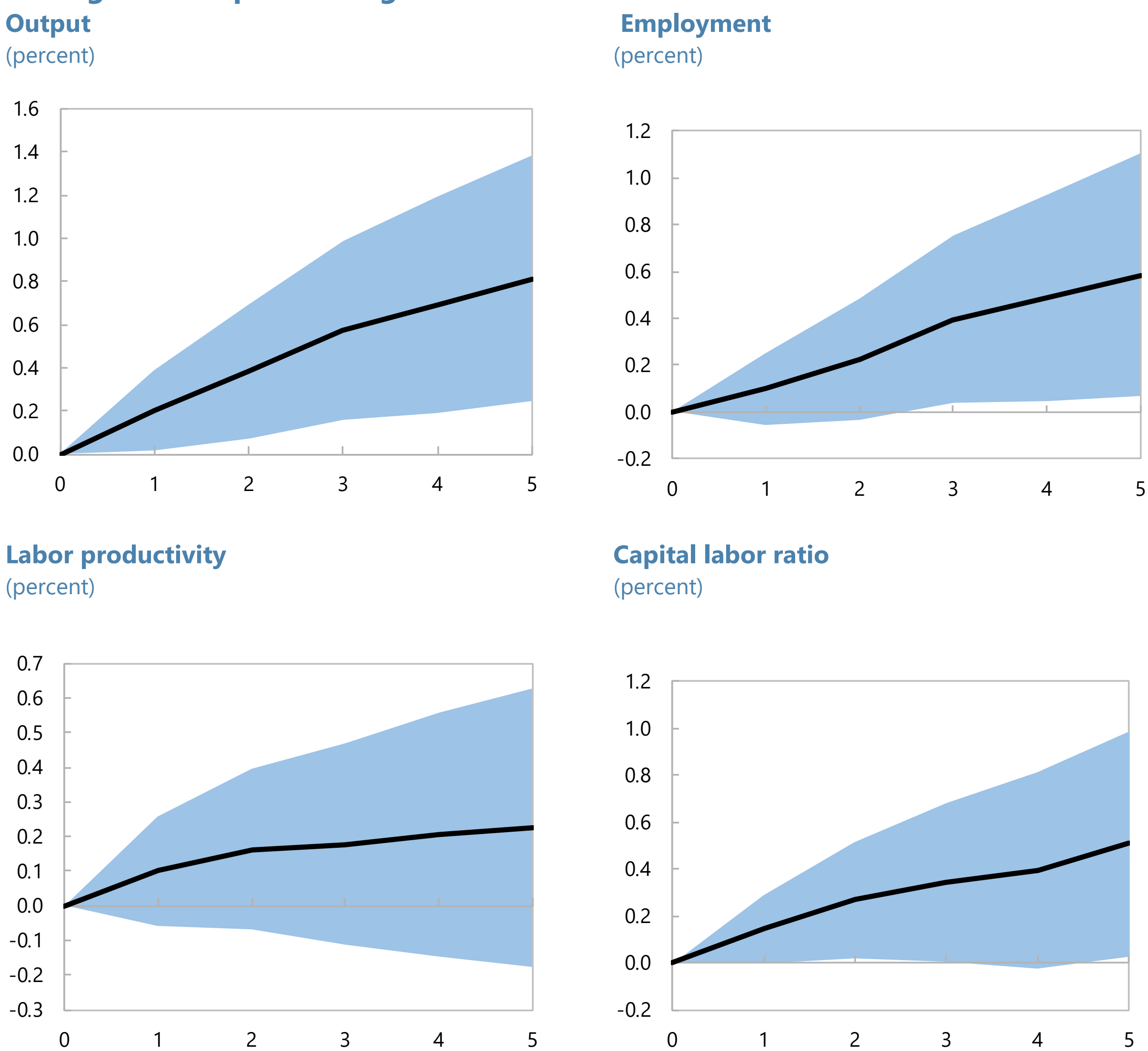
31. The coefficient of interest is β_1^h, which captures the cumulative impact of the time *t* migration shock at horizon *h*. Various robustness exercises are conducted which include additional control variables—lags of the dependent variable, real GDP growth rate, and lagged native employment rate growth. Despite efforts to create an exogenous instrument, its performance based on F-statistics, and controlling for other possible omitted variables it may still be the case that there is some endogeneity in our estimates. This could be the case, for example, if there is a prolonged (over 10 years) upswing in the economic conditions in the destination countries that act as a pull factor and thereby makes the shift-share IV not perfectly exogenous. Given these constraints, various robustness specifications are estimated to control for omitted variable bias, but one should still interpret the coefficient estimates as an upper bound for the impact on macroeconomic variables.

32. Results for the estimation of equation (1) are reported in the impulse response functions in the chapter's main text. Several robustness checks are conducted. First, IMF *World Economic Outlook* data is used for employment, output, and labor productivity. Results for these impulse response function, as well as the capital-labor ratio are reported in Figure 1. The results are broadly robust, except the positive impact on employment is now statistically significant three years after the shock. This may be due to the definitions of employment, which in the *World Economic Outlook* database include only those formally employed while in the Penn World Table includes all *engaged*, that is, also include self-employed or in the informal sector.

33. Additional robustness specifications, which control for possible omitted variable bias are also conducted. These include controlling for lags of the dependent variable, for lagged real GDP growth, and for lagged native employment growth. All results are broadly robust. Table 1 reports results for the first robustness exercise (lagged dependent variable), and Table 2 for a specification that includes all control variables (lagged dependent variable, lagged real GDP growth, lagged native employment growth). These tables show the coefficient estimates for the specification across all horizons.

34. Finally, Table 3 reports the estimate impact of refugee shocks on the macroeconomic variables of interest, based on the specification defined in (1). Unlike migration shocks, refugee shocks have a much less significant impact on the macroeconomy of their host country. This is consistent with the fact that refugees arrive in much smaller numbers and are often constrained to camps or certain geographical areas without the opportunity to participate in the local economy.

Figure 1. Impact of Migration Shock – IMF *World Economic Outlook* Data



Source: IMF staff estimates.
Notes: This panel depicts the effect of a 1 percent increase in the migration inflow to employment ratio in the destination country on the macroeconomic variables indicated, estimated on a sample of OECD countries from 1980-2018 using the local projections method of Jorda (2005). Horizon 0 is the year prior to the shock, horizon 1 shows the effect of the shock on impact. The solid line represents the impulse response estimates and the shaded areas represent 90 percent confidence intervals.

Table 1. Impact of Migration Shock on Macroeconomic Variables: Control for lagged dependent variable

Horizon:	1	2	3	4	5	6
Output	0.100	0.304	0.491*	0.803**	0.870**	0.783*
	(0.112)	(0.187)	(0.263)	(0.342)	(0.403)	(0.445)
Employment	-0.00155	0.0358	0.120	0.209	0.224	0.180
	(0.0555)	(0.115)	(0.180)	(0.248)	(0.301)	(0.339)
Labor Productivity	0.0897	0.220*	0.306*	0.506***	0.550**	0.524**
	(0.0895)	(0.133)	(0.166)	(0.193)	(0.213)	(0.235)
TFP	0.0428	0.120	0.196	0.459**	0.559**	0.595**
	(0.0990)	(0.144)	(0.180)	(0.208)	(0.226)	(0.245)
Capital employment ratio	0.123**	0.295***	0.483***	0.653***	0.842***	1.022***
	(0.0531)	(0.102)	(0.153)	(0.204)	(0.239)	(0.266)
Unemployment Rate	0.0735**	0.167**	0.269**	0.345**	0.362**	0.426**
	(0.0370)	(0.0773)	(0.115)	(0.151)	(0.179)	(0.186)
Native employment	0.0799	0.0855	0.0303	-0.169	-0.355	-0.146
	-0.0772	-0.134	-0.208	-0.281	-0.32	-0.339
Native unemployment rate	0.0610	0.116	0.211	0.300	0.248	0.333
	(0.0591)	(0.117)	(0.169)	(0.225)	(0.267)	(0.259)
Control Variables	Lagged dependent	Lagged dependent	Lagged dependent	Lagged dependent	Lagged dependent	Lagged dependent
N[1]	782	750	718	686	654	622
Sample	OECD	OECD	OECD	OECD	OECD	OECD
Country FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes

1 Sample for native employment and unemployment is 382 and 335, respectively, at H=1.

Source: IMF staff estimates.

Note: All dependent variables, except the unemployment rate, are expressed in logs.

Table 2. Impact of Migration Shock on Macroeconomic Variables: Control for lagged dependent variable, real GDP, and native employment growth

Horizon:	1	2	3	4	5	6
Output	0.168	0.359	0.448	0.920**	1.237**	1.259**
	(0.138)	(0.220)	(0.330)	(0.443)	(0.500)	(0.496)
Employment	0.00302	-0.0448	-0.0450	0.112	0.390	0.330
	(0.0767)	(0.140)	(0.226)	(0.324)	(0.384)	(0.382)
Labor Productivity	0.189	0.455**	0.611**	0.930***	1.021***	1.117***
	(0.130)	(0.190)	(0.257)	(0.319)	(0.342)	(0.365)
TFP	0.164	0.388**	0.554**	1.083***	1.191***	1.158***
	(0.128)	(0.186)	(0.251)	(0.320)	(0.362)	(0.365)
Capital employment ratio	0.0971	0.291**	0.435**	0.389	0.293	0.459
	(0.0676)	(0.123)	(0.189)	(0.254)	(0.289)	(0.291)
Unemployment Rate	0.108**	0.224**	0.312*	0.374	0.232	0.177
	(0.0507)	(0.108)	(0.169)	(0.239)	(0.290)	(0.292)
Native employment	-0.0587	-0.0681	0.0825	0.0827	0.178	0.518
	(0.0880)	(0.181)	(0.313)	(0.448)	(0.522)	(0.517)
Native unemployment rate	0.101*	0.190*	0.259	0.320	0.173	0.143
	(0.0582)	(0.113)	(0.165)	(0.227)	(0.270)	(0.270)
Control Variables	Lagged native employment growth; Lagged dependent; Lagged RGDP	Lagged native employment growth; Lagged dependent; Lagged RGDP	Lagged native employment growth; Lagged dependent; Lagged RGDP	Lagged native employment growth; Lagged dependent; Lagged RGDP	Lagged native employment growth; Lagged dependent; Lagged RGDP	Lagged native employment growth; Lagged dependent; Lagged RGDP
N	359	359	330	302	276	248
Sample	OECD	OECD	OECD	OECD	OECD	OECD
Country FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes

Source: IMF staff estimates.

Note: All dependent variables, except the unemployment rate, are expressed in logs.

Table 3. Impact of Refugee Shock on Macroeconomic Variables:

Horizon:	1	2	3	4	5	6
Output	0.252	0.324	0.172	0.160	0.453	0.291
	(0.299)	(0.433)	(0.531)	(0.628)	(0.641)	(0.628)
Employment	0.223	0.441*	0.761***	0.803***	0.360	-0.0462
	(0.158)	(0.266)	(0.252)	(0.280)	(0.477)	(0.419)
Labor Productivity	0.667	1.263	0.810	0.598	1.509	2.457
	(0.889)	(1.189)	(1.229)	(1.268)	(1.572)	(2.071)
TFP	0.793	0.827	0.135	0.125	1.371	1.391
	(0.686)	(0.765)	(0.903)	(1.194)	(1.474)	(1.152)
Capital employment ratio	-0.169	-0.289	-0.565	-0.626	-0.129	-0.113
	(0.169)	(0.315)	(0.391)	(0.473)	(0.680)	(0.661)
Unemployment Rate	0.168	0.466**	0.439	0.129	-0.0194	-0.820**
	(0.105)	(0.222)	(0.351)	(0.298)	(0.456)	(0.338)
Control Variables	N/A	N/A	N/A	N/A	N/A	N/A
N	3052	2949	2846	2742	2638	2534
Sample	EMDE	EMDE	EMDE	EMDE	EMDE	EMDE
Country FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes

Source: IMF staff estimates.

Note: All dependent variables, except the unemployment rate, are expressed in logs.

E. Role of Policy

In order to study how policy influences the macroeconomic impacts of immigration, the specification in (1) is augmented by adding interactions between the immigration shock and various policy indicators or macroeconomic conditions. Specifically:

$$y_{i,t+h} - y_{i,t-1} = \alpha_i^h + \gamma_t^h + \beta_1^h \frac{\Delta Mig_{it}}{E_{i,t-1}} + \beta_2^h \frac{\Delta Mig_{it}}{E_{i,t-1}} \cdot Policy_{it} + \beta_3^h Policy_{it} + \varepsilon_{it}^h \qquad (2)$$

Where $Policy_{it}$ is, in turn: trade, product, and labor sector reform index, financial sector reform index, active labor market policy spending per capita (lagged 5 years to allow for policies to have an impact), spending on post-secondary non-tertiary education (as a share of GDP, lagged 5 years), indices of immigration integration restrictions and immigration border and internal controls (where a higher value of the index indicates greater control), and an index of exposure to routinization. Due to data limitations, the role of policies is not examined for the case of refugee inflows.

Equation (2) is estimated via two-stage least squares, as in equation (1), but the policy variable is not instrumented. As a result, the coefficients should be interpreted strictly as correlations as there may be reverse causality between macroeconomic outcomes and migration policies. As a result, the coefficient estimates will be larger than they would be if their causal impact could be identified.

35. The main coefficient of interest in equation (2) is β_2^h which indicates whether the impact of a migration shock has differential impacts depending on the level of a given policy in the host country. The impulse response function for the linear combination of $\hat{\beta}_1^h + \hat{\beta}_2^h Policy_{it}$, which represents the total impact of a migration shock for a given level of the policy variable is also of interest. Figure 20 in the main text shows the impact on employment for the mean plus and minus two standard deviation of various policy measures. Table 4 report the corresponding coefficient estimates.

Table 4. Impact of Migration Shock on Employment, by Policy

Horizon:	1	2	3	4	5
Migration Shock	0.0884	0.147	0.212	0.255	0.192
	(0.0777)	(0.151)	(0.228)	(0.304)	(0.372)
Migration Shock*ALMP (lag 5 years)	-5.397	-2.008	19.63	50.22	88.17*
	(11.40)	(22.00)	(31.32)	(37.51)	(47.98)
ALMP (lag 5 years)	-0.0642	-0.256	-0.814	-1.355*	-1.760**
	(0.262)	(0.461)	(0.648)	(0.725)	(0.724)
Migration Shock	0.240*	0.209	0.0935	0.0192	-0.208
	(0.139)	(0.269)	(0.410)	(0.635)	(0.788)
Migration Shock*Education Spending (lag 5 years)	0.0844	0.494	1.085	1.568	2.152*
	(0.222)	(0.432)	(0.671)	(1.004)	(1.251)
Education Spending (lag 5 years)	-0.00757	-0.0144	-0.0107	0.000956	0.0248
	(0.00922)	(0.0166)	(0.0252)	(0.0302)	(0.0371)
Migration Shock	0.628*	1.104*	1.027	0.526	-0.0317
	(0.335)	(0.599)	(0.824)	(1.026)	(1.297)
Migration Shock* Real Sector Reforms	-0.642	-1.084	-0.865	-0.199	0.532
	(0.437)	(0.787)	(1.079)	(1.336)	(1.673)
Real Sector Reforms	-0.000444	0.0568	0.0944*	0.135**	0.138**
	(0.0310)	(0.0474)	(0.0564)	(0.0611)	(0.0695)
Migration Shock	-0.787	-1.694	-3.081	-4.307	-6.576*
	(1.007)	(1.813)	(2.507)	(3.140)	(3.735)
Migration Shock* Financial Sector Reforms	1.005	2.102	3.671	4.972	7.442*
	(1.083)	(1.940)	(2.676)	(3.335)	(3.958)
Financial Sector Reforms	0.0197*	0.0178	0.00558	-0.0160	-0.0363
	(0.0115)	(0.0211)	(0.0292)	(0.0347)	(0.0389)
Migration Shock	0.438**	0.788**	1.124**	1.222**	1.324*
	(0.184)	(0.329)	(0.474)	(0.601)	(0.690)
Migration Shock*Migration Integration Restrictivene	-1.152*	-1.998*	-2.810*	-3.165	-3.686
	(0.590)	(1.123)	(1.682)	(2.178)	(2.502)
Migration Integration Restrictiveness	0.0325*	0.0664**	0.0889**	0.115**	0.138**
	(0.0174)	(0.0315)	(0.0438)	(0.0530)	(0.0605)
Migration Shock	0.319*	0.594*	0.811	1.513**	1.965**
	(0.191)	(0.352)	(0.508)	(0.746)	(0.959)
Migration Shock*Routinization index	-0.202	-1.512	-2.985	-7.344*	-10.20*
	(1.167)	(2.096)	(3.030)	(4.335)	(5.635)
Routinization index	-0.0238	-0.0563	-0.109*	-0.142*	-0.203*
	(0.0326)	(0.0493)	(0.0647)	(0.0838)	(0.107)
Control Variables	N/A	N/A	N/A	N/A	N/A
Sample	OECD	OECD	OECD	OECD	OECD
Country FE	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes

Annex 4.4: Model Simulations

A. Model Description

36. The simulations are based on the G20MOD, a module of the IMF's Flexible System of Global Models (FSGM). G20MOD is an annual, multi-region, forward-looking, model of the global economy combining both micro-founded and reduced-form formulations of economic sectors. G20MOD contains individual blocks for the G-20 countries, and 5 additional regions to cover the remaining countries in the world. Each country/regional block is largely distinguished by its unique parameterization. Each economy in the model is structurally identical, but with different key steady-state ratios and behavioral parameters. A detailed description of FSGM and its simulation properties can be model is presented in greater detail in Andrle et al. (2015).

37. In the model, output is produced using capital and labor, with labor provided by households. Investment is driven by decisions of profit-maximizing and forward-looking firms, subject to investment adjustment costs, resulting in a version of the Tobin's Q model. The cost of borrowing of firms is affected by an endogenous risk premium, which increases in a downturn (financial accelerator). Labor is provided by households at the market wage, with household choosing the rate of labor-force participation.

38. Private consumption in G20MOD is driven by two types of households, optimizing overlapping-generations (OLG) households with access to financial markets, and liquidity-constrained households (LIQ). LIQ households consume their full disposable income every period. Their disposable income is formed by their after-tax labor income, transfers from the government, and received remittances, if applicable.

39. Imports of goods and services are driven by the relative prices of domestic and foreign goods and import requirements of domestic consumption, investment, government expenditures, and exports. Exports of goods and services are given by trading partners demand for imports.

40. There is a fully specified fiscal sector with several policy instruments. There is full stock-flow accounting of fiscal policy with a fiscal balance that accumulates to a stock of debt and tracking of the interest cost. Fiscal policy stabilizes debt as a percent of GDP in the long-run and responds to the output gap in the short-term to enhance macro stability. On the revenue side, households are subject to a labor-income tax, an ad-valorem consumption tax, and lump-sum taxes. Firms pay capital income taxes. On the expenditure side, the model features government consumption, productive government investment, transfers to households, and the interest cost of the outstanding government debt.

41. Monetary policy is represented by an interest rate reaction function where the standard form is an inflation-forecast-based rule operating under a flexible exchange rate. However, this can also be modeled as a fixed exchange rate (currency union) or managed floating exchange rate.

42. There are three commodities in the model—oil, metals, and food. The consumer price index excluding food and energy is determined by an inflation Phillips curve. Prices are sticky and reflect the expected paths of import prices and the economic cycle, as captured by the output gap. Wage inflation exhibits stickiness and allows the real wage to return to its equilibrium only gradually depending on the expected evolution of overall economic activity.

43. G20MOD captures a complete set of bilateral migration and remittance flows. The population, labor force, and employment distinguish between "domestic" and "foreign" households. The wage differential between the two groups is used as a proxy for the relative productivity of the two groups. The relative productivity of both types of labor are reflected in the potential output of the economy. Expatriate workers are assumed to remit a fraction of their disposable income to their countries of origin. All expatriate workers are assumed to be liquidity-constrained, consuming all their disposable income left after sending the remittances. Details on the calibration of remittances and of immigration shares are found in Snudden (2017).

B. Calibration

44. The table below presents the list of countries/groups for which immigration flows were simulated. The calibration of the labor market outcomes of immigrants in these recipient countries is based on OECD (2018).[5] The assumption that the wage gap of immigrants closes in about 15 years is consistent with National Academies of Sciences, Engineering, and Medicine (2017).

Destination	Employment rate (percent)	Initial wage gap (in percentage of natives' wage)
Germany	67	87
France	57	85
Italy	60	67
Spain	60	65
Rest of euro area	61	75
USA	70	83
Russia	69	92
Saudi Arabia	76	92
Remaining oil exporters	69	92

[5] Participation rates of immigrants for Russia and for the Remaining oil exporters were missing and have been replaced with the average values of the remaining OECD countries. The same holds for the initial wage gap in Russia, Saudi Arabia and Remaining oil exporters.

Annex 4.5: Labor Migration, Trade, and Productivity Growth in ASEAN[1/]

The Association of South East Asian Nations (ASEAN) countries have witnessed large within-region migration flows in recent years driven by employment opportunities arising from within-region income and skill differentials. While migrant workers have contributed to economic growth in host countries, the latter have also expressed concern that an overreliance on low-skilled migrants is impeding a move up the value chain. Recent surveys also reveal that host countries typically worry that low-skilled migrants adversely affect employment and wages of local workers (ILO 2012) and support strict limits on immigration policy (World Values Survey 2015).

Using a spatial dynamic general equilibrium model for trade and labor, this box explores the effect of a hypothetical tightening in immigration policy on the cross-border flows of low- and high-skilled workers, and the impacts of these worker flows on employment, GDP, and productivity growth across sectors and in both host and sending countries. In an illustrative scenario, Malaysia, a major migration recipient country in the region, is assumed to impose a strict ban on the inflows of low-skilled migrant workers while high-skilled migrants can move freely across the border. The results of the simulations are similar if the restriction on migration is imposed by other migration hub countries, such as Thailand.[2/]

Effects on the Host Country

The ban creates a shortage of low-skilled workers in the host country, pushing up their average real wages by 8 percent over the next 10 years, hence raising unit production costs. The resulting loss in trade competitiveness reduces demand for the host country's intermediate and final goods in both domestic and international markets; this slows the inflow of high-skilled migrants and raises the unemployment rate of local high-skilled workers. The resulting decrease in the wages of the high skilled further discourages skilled workers to migrate to the host country in subsequent periods. A moderation in both low- and high-skilled employment growth, in turn, implies an average annual GDP loss of about 1 percent. Even under the extreme assumption of perfect substitutability between native and immigrant workers, the ban leads to only a marginal improvement in the host country's productivity growth (0.1 percent) as a small share of high-skilled workers shifts from the production of goods to research and development jobs. The expected technology upgrade will not happen automatically because competitive firms will not choose to expand investment in research and development, given its high-risk nature and positive spillovers.

Heterogeneous Impacts and Spillovers

The ban has a larger impact on high-skilled employment in industries that are more integrated in global value chains, such as the electrical and electronic industry for Malaysia. While such a ban, if implemented in Malaysia, could also adversely affect growth in countries belonging to the same supply chain or that export heavily to Malaysia, it could also raise growth and productivity in countries that would absorb the outflow of high-skilled workers from Malaysia.

1/ The author of this box is Xin Li.

2/ The model extends Caliendo and others (2019) by introducing households of different skill levels that make migration decisions based on country and sector-specific wage differentials and migration cost. The model takes into account the buildup of knowledge capital and abstracts from physical capital adjustment. The model was calibrated to the ASEAN-5, Vietnam and major trading partners. The model is stylized in that it assumes perfect substitutability between immigrants and natives. In this sense, the impacts on GDP from the immigration ban can be interpreted as upper bounds.

References

Alesina, Alberto, Davide Furceri, Jonathan Ostry, Chris Papageorgiou, and Dennis Quinn. 2020. F "Structural Reforms and Electoral Outcomes: Evidence from a New Database of Regulatory Stances and Policy Changes." NBER Working Paper No. 26720, National Bureau of Economic Research, Cambridge, MA.

Andrle, Michal, Patrick Blagrave, Pedro Espaillat, Keiko Honjo, Benjamin Hunt, Mika Kortelainen, René Lalonde, Douglas Laxton, Eleonora Mavroeidi, Dirk Muir, Susanna Mursula, and Stephen Snudden. 2015. "The Flexible System of Global Models – FSGM." IMF Working Paper 15/64, International Monetary Fund, Washington, DC.

Beine, Michel, and Christopher Parsons. 2015. "Climatic Factors as Determinants of International Migration." *The Scandinavian Journal of Economics* 117 (2): 723–67.

Beine, Michel, Simone Bertoli, and Jesús Fernández-Huertas Moraga. 2016. "A Practitioners' Guide to Gravity Models of International Migration." *The World Economy* 39 (4): 496–512.

Card, David. 2009. "Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration." *Journal of Labor Economics* 19 (1): 22–64.

Cameron, A. Colin, and Pravin K. Trivedi. 2005. *Microeconometrics: Methods and Applications.* Cambridge, UK: Cambridge University Press.

Cattaneo, Cristina, and Giovanni Peri. 2016. "The Migration Response to Increasing Temperatures." *Journal of Development Economics* 122: 127–46.

Clark, Ximena, Timothy J. Hatton, and Jeffrey G. Williamson. 2002. "Where Do US Immigrants Come From, and Why?" NBER Working Paper No. 8998, National Bureau of Economic Research, Cambridge, MA.

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.

Hatton, Timothy J., and Jeffrey G. Williamson. 2003. "Demographic and Economic Pressure on Emigration Out of Africa." *Scandinavian Journal of Economics* 105 (3): 465–86.

Hausman, Jerry A., Bronwyn H. Hall, and Zvi Griliches. 1984. "Econometric Models for Count Data with an Application to the Patents-R&D Relationship." *Econometrica* 524: 909–38.

Helbling, Marc, Liv Bjerre, Friederike Römer, and Malisa Zobel. 2017. "Measuring Immigration Policies: The IMPIC Database." *European Political Science* 16: 79–98.

Jordà, Òscar. 2005. "Estimation and Inference of Impulse Responses by Local Projections." *American Economic Review* 95 (1): 161–82.

Laeven, Luc, and Fabian Valencia. 2013. "Systemic Banking Crises Database." *IMF Economic Review* 61 (2): 225–70.

National Academies of Sciences, Engineering, and Medicine. 2017. *The Economic and Fiscal Consequences of Immigration.* Washington, DC: The National Academies Press.

Organisation for Economic Co-operation and Development (OECD). 2018. *Settling in 2018, Indicators of Immigrants' Integration*. Paris: OECD.

Midgley. 2018. *Groundswell: Preparing for Internal Climate Migration.* Washington, DC: World Bank.

Sarvimäki, Matti, and Kari Hämäläinen. 2016. "Integrating Immigrants: The Impact of Restructuring Active Labor Market Programs." *Journal of Labor Economics* 34 (2): 479–508.

Schmid, Samuel, and Marc Helbling, 2016. "Validating the Immigration Policies in Comparison (IMPIC) Data Set." Discussion Paper SP VI 2016–202, WZB Berlin Social Science Center.

Snudden, Stephen. 2017. "International Remittances, Migration, and Primary Commodities in FSGM." IMF Working Paper 17/20, International Monetary Fund, Washington, DC.

White, Halbert. 1980. "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity." *Econometrica* 48 (4): 817–88.