# Appendix A Robustness of the migration policy selectivity indexes

In this section, we check the sensitivity of our indexes of selectivity to our modelling choices. We specifically focus on the main assumption, namely that our list of legislative changes is complete and without errors. The following robustness checks by no means constitute a formal test of these assumptions, but instead gauge the sensitivity of the selectivity indexes to the purposeful introduction of errors.

## A.1 Scaling errors

The first robustness check looks at the impact of the DEMIG encoded scale of legislative changes. Each legislation is assigned a magnitude score between 1 and 4 by **?** categorising them as either fine-tuning, minor, mid-level or major change. In order to see how errors in this scoring affect the selectivity indexes, we gave all changes the same magnitude score of one, leaving only the direction of the effect (increase or decrease in restrictiveness) intact. However, this barely impacted the indexes, as can be seen from Table A-1. The correlations between the baseline results and the robustness checks is in excess of 96% for all indexes, regardless of whether we looked at the overall correlation or only considered the correlation between counties (between) or over time (within).

**Table A-1:** Correlation with the baseline index values

| | Overall | Between | Within |
|---|---|---|---|
| $MPS^{skill}$ | 0.96 | 0.96 | 0.96 |
| $MPS^{res}$ | 0.97 | 0.98 | 0.97 |
| $MPS^{nat}$ | 0.99 | 1.00 | 0.98 |

## A.2 Initial anchor values

The third robustness verifies the importance of taking the year 1945 as the initial anchor value of zero given that the initial value of policy selectivity is unknown. To start with, we changed the anchor point from 1945 to 1960 to see how this impacts the indicator values. In both the baseline and the alternative scenario, the start date of the dataset was kept in 1990, meaning that the alternative scenarios reduced the burn-in period from 45 to 30 and even 10 years. As can be seen in Table A-2, the effect on the indicator values was minimal.

**Table A-2:** Correlation with the baseline index values

| startyear | 1960 | 1980 |
|---|---|---|
| $MPS^{res}_{i,t}$ | 0.9937 | 0.9752 |
| $MPS^{skill}_{i,t}$ | 0.9740 | 0.8470 |
| $MPS^{nat}_{ij,t}$ | 0.9539 | 0.8484 |

There are a number of reasons why the choice of anchor point has a modest impact on the resulting indicator. To start, the number of legislative changes per year listed in the DEMIG database is heavily skewed towards the latter years. only about 35% of changes happens in the 45 years between 1945 and 1989, while the next 25 years hold the remaining 65%. Furthermore, the effect of a single legislative change is also fairly limited, with the maximum impact capped at a plus or minus four. Finally, there is also a strong positive correlation between the legislative changes over time. For example, countries that strongly increase the restrictiveness in one year are more likely to continue doing so. For the starting values to have a significant distorting effect, you would need some countries to radically alter the restrictiveness of their policies from one year to the next.

## A.3 Policy decay

The final robustness check is focused specifically on our use of a running sum. As was explained above, using a running sum means that all errors in the dataset are compounded. As a result, the longer the running sum, the larger the uncertainty of the estimates becomes. As an alternative modelling choice, we replace our computation of the level of migration policy from $L_{i,t} = L_{i,t-1} + C_{i,t}$ to one where the effects fade over time: $L_{i,t} = \delta L_{i,-1} + C_{i,t}$. Lacking any information on the speed with which policy should decay, we set the value of $\delta$ such that the effect of a legislative change is reduced to only 5% after 20 years. This gave us a relatively short half-life of 4.6 years, which serves as a lower bound as the actual persistence of policy is likely much higher. It should also be noted that the DEMIG dataset does contain information on the role-back of legislation, which conflicts with the fade-out.

The effect of allowing policy to fade out over time is slightly larger than that of our first two robustness checks. However, the correlation with our baseline results remains high as can be seen form Table A-3. Overall, our selectivity indexes seem to be highly robust to our modelling choices as well as potential errors or omission in the dataset.

**Table A-3:** Correlation with the baseline index values

| | Overall | Between | Within |
|---|---|---|---|
| $MPS^{skill}_{i,t}$ | 0.71 | 0.83 | 0.74 |
| $MPS^{res}_{i,t}$ | 0.81 | 0.95 | 0.80 |
| $MPS^{nat}_{ij,t}$ | 0.87 | 0.93 | 0.80 |

# Appendix B List of countries by sub-groups

**OECD**

**EU**

| | |
|---|---|
| Austria | Hungary |
| Belgium | Ireland |
| Czech Rep. | Italy |
| Germany | Luxembourg |
| Denmark | Netherlands |
| Spain | Poland |
| Finland | Portugal |
| France | Slovakia |
| UK | Slovenia |
| Greece | Sweden |

**non-EU**

| | |
|---|---|
| Australia | South Korea |
| Canada | Mexico |
| Switzerland | Norway |
| Chile | New Zealand |
| Iceland | Turkey |
| Israel | USA |
| Japan | |

**non-OECD**

Argentina
China
Brazil
Indonesia
India
Morocco
Russia
Ukraine
South Africa

# Appendix C Data sources for the explanatory variables

| **Variable** | **Source** |
|---|---|
| Interdecile earnings ratio (P90P50) | Labour force Statistics (OECD) |
| GDP per capita (PPP) | World Development Indicators (World Bank) |
| Bilateral migrant stocks | International Migration Database (OECD) |
| Unemployment | World Indicators of Skills for Employment OECD) |
| Migration policy restrictiveness | **?** |
| Migration policy selectivity | Own computation |

# Appendix D Imputing the DIOC education and occupation migration data

As explained in the main body of the paper, our regression models require data on the number and share of high-skilled migrants and managers and businesspeople migrating each year. However, the DIOC data we have at our disposal cover only the stock of migrants at a rather low frequency (in 2000, 2005 and 2010) and with a large fraction of country-pairs missing the intermediate measurement. To obtain net migration flows, we take the 5-yearly differences in the stock data, which causes the range of destination countries to fall substantially. While it is possible to run the regression with this dataset, it risks providing a selective view of the impact of the new indicators of migration selectivity.

For this reason, we use a statistical model to fill in the gaps in the 2005 DIOC data. Following **?**, we construct a (Bayesian) state-space model consisting of two main sets of equations. The state equations describe the dynamic behaviour and relationships between our main variables of interest (the latent state variables). In this case, our state variables are the stock and flow of high-skilled migrants in the first model and the stock and flow of managers and businesspeople in the second. The relationship between the stock and flows is described in a demographic model. The second set of equations is the measurement equations, which describe the relationship between these state variables and our observed data, e.g., how the flow of high-skilled migrants relates to the overall flow of migrants. See **?** or **?** for more information on state-space models.

Using a state-space model allows us to combine data on migration stocks and flows of different sources with a demographic model to help estimate the most likely value for our missing data. This technique is related to some extent to the demographic accounting technique employed by (**??**) to impute net migration flow data, in which differences in stock data are combined with demographic data. The demographic accounting technique requires close-to-complete information on the stock of migrants. These are used to build contingency tables that describe where the population of a particular origin country is distributed around the world. Given the paucity of data on managers or high-skilled, we instead use the state-space approach, allowing us to estimate the missing data for each country pair separately.

Note that the imputation algorithm would in principle allow us to use yearly imputed values of migration. However, the estimation results using the yearly data differ considerably from those using the original DIOC data. As we could not rule out that these differences were caused by the imputation algorithm (rather than e.g. sample selection effects), we have limited our statistical analysis to the 5-yearly change in migration flows used in the DIOC data.

## D.1 Imputation algorithm

We used two different models to impute both the occupation and education DIOC data. However, as both models are very similar, we will focus our explanation on the imputation of high-skilled migrants. The only difference between both regressions occurs in the measurement equation (A-5), where the share of high-skilled migrants in the origin country is replaced by the share of managers in the origin country.

The state equation is built on a demographic identity: the only way in which the stock of migrants based on country of birth can change is if migrants enter the country, leave the country or die.[1] If $S^h_{ij,t}$ is the stock of high-skilled migrants from $i$ in $j$ at time $t$, $N^h_{ij,t}$ are the net flows from $i$ to $j$ and $D^h_{ij,t}$ is number of high-skilled migrants from $i$ in $j$ that have died in year $t$, this gives us the following equation:

$$S^h_{ij,t} \equiv S^h_{ij,t-1} + N^h_{ij,t} - D^h_{ij,t} \tag{A-1}$$

For the vast majority of countries, the information on how many migrants have died per origin country is not available. We follow the approach of **?** and assume that the deaths are equal to the stock of migrants already in the country multiplied by a destination-country-specific death rate.

$$D^h_{ij,t} = \delta_{i,t}\, S^h_{ij,t-1}, \tag{A-2}$$

As many of the variables that influence the flow of migration are highly persistent (e.g., size of the migrant population, population size of the sending country), we also want to allow for this persistence in the net migration flows. To that end, we model this variable as an autoregressive process with one lag process. The level of persistence in these flows is estimated within the model.

$$\begin{aligned} N^h_{ij,t} &= \tau_{ij}\, N^h_{ij,t-1} + \mu_{ij,t} \\ \mu_{ij,t} &\sim N(0, \sigma^\mu_{ij}) \end{aligned} \tag{A-3}$$

The measurement equation consists of two parts. To anchor our results, we impose that the available migration stock data from DIOC is correct.

$$DIOC_{ij,t} = S^h_{ij,t} \tag{A-4}$$

[1]Depending on the legal system, babies born from migrant mothers are counted as an increase in the domestic population, or as an increase in the net migration flow. Either way, the births are already taken into account.

The second equation relates the flow of high-skilled migrants to the total migration flow and the share of high-skilled individuals in the origin country. If the choice to migrate was independent of skill level, then multiplying both variables would provide a good approximation of the flow of high-skilled migrants. However, as skill level is likely to influence the likelihood (and ability) to migrate, we embed this relationship in a linear error model.

$$\begin{aligned} hs_{j,t} * N_{ij,t} &= zN^h_{ij,t} + c_{ij} + \epsilon_{ij,t} \\ \epsilon_{ij,t} &\sim N(0, \sigma^{\epsilon}_{ij}) \end{aligned} \tag{A-5}$$

$c_{ij}$ and $z$ capture the persistent differences between the flow of high-skilled migrants, $N^h$, and the error term $\epsilon$ accounts for any stochastic deviations. As the magnitude of the flow and stock of migrants can be very different depending on the countries in question, the constant $c_{ij}$ and variance of the error term $\sigma^{\epsilon}_{ij}$ can differ for each country-pair.

Putting these equations together, we get the following state-space model:

$$\begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} S^h_{ij,t} \\ N^h_{ij,t} \end{bmatrix} = \begin{bmatrix} 1-\delta_{j,t} & 0 \\ 0 & \tau^N \end{bmatrix} \begin{bmatrix} S^h_{ij,t-1} \\ N^h_{ij,t-1} \end{bmatrix} + \begin{bmatrix} 0 \\ \mu_{ij,t} \end{bmatrix} \tag{A-6}$$

$$\begin{bmatrix} DIOC_{ij,t} \\ hs_{ij,t} N_{ij,t} \end{bmatrix} = \begin{bmatrix} 0 \\ c_{ij} \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & Z \end{bmatrix} \begin{bmatrix} S^h_{ij,t} \\ N^h_{ij,t} \end{bmatrix} + \begin{bmatrix} 0 \\ \epsilon_{ij,t} \end{bmatrix} \tag{A-7}$$

## D.2 Data sources

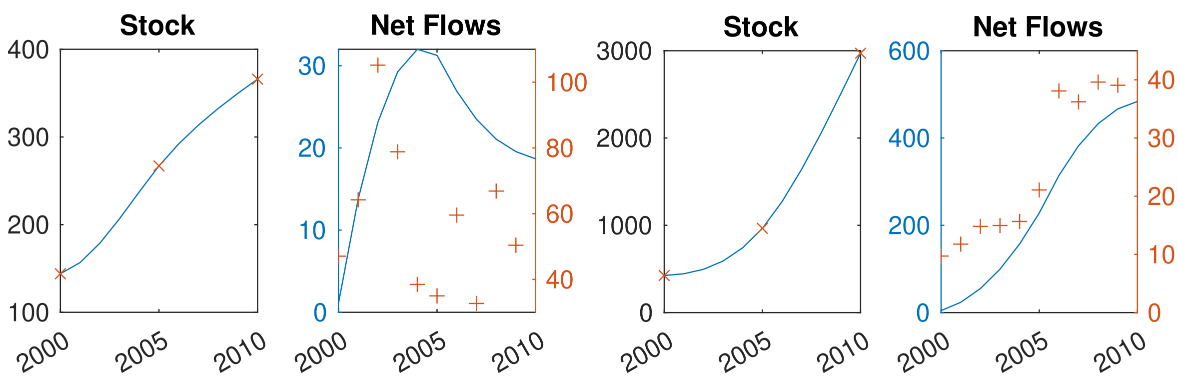
In addition to the DIOC dataset, several other databases were used in the computations. We collected information on the yearly flow of migrants from the OECD's International Migration Database. To proxy the inflow of managers, we multiplied the total inflows with the share of managers in the total population of the origin country (proxied by the ISCO-1 occupation category, using population and labor force data from the ILO (**?**).[2] Unfortunately, a similar indicator for education was harder to come by. While the World Bank has a variable measuring the share of highly educated people in the total population, this variable is missing for most of the dataset. As a result, we used the average number of years of schooling from the UNDP Human Development Report 2020 instead. Finally, death rates of the destination countries from the WHO's Global Health Observatory.

[2]See for concepts, definitions and a description of the methodology https://ilostat.ilo.org/resources/concepts-and-definitions/description-employment-by-occupation

## D.3 Results

The estimation model ran for 5,000 iterations, of which the first 4,000 were discarded as burn in.[3] The remaining iterations were used to compute the most likely bilateral stock and net flows of high-skilled migrants and managers and businesspeople. In this way, the data set for migration according to education increases to more than $37,000$ observations (resulting in a sample for the estimations of some $8,000$ observations ranging over 27 destination countries, after adding the control variables) and approximately $23,000$ observations for occupation (giving a final data set of $6,300$ observations for 25 destination countries).

Figure 1 compares the imputed values of the occupation and education data to the source data. In both cases we see that the DIOC stock data anchors the imputed stock values (left panel), while the flows try to follow the pattern in the $hs_{j,t} * N_{ij,t}$ variable (right panel). However, this is not the case for all country-pairs. For example, according to the DIOC data, no high-skilled migrants were migrating from Poland to Chile in 2005 or 2010. As a result, the model returns all zeros for the intervening years as well.




**(a)** High-skilled migration from Israel to New Zealand **(b)** Managers migrating from Congo, PDR to France

**Figure 1:** Comparison of the imputed data and source data
Comparison of the imputed migration level (blue lines) with the source data (red crosses). The left-hand panel shows the stock data and DIOC data, while the right-hand panel shows the net flows and our proxy for the inflow of high-skilled migrants.

[3]We used uninformative priors and checked the model's convergence using a visual inspection of the parameters plots, autocorrelation function and CUMSUM graphs.

# Appendix E Robustness regressions

**Table A-4:** Origin-time and origin-destination fixed effects

| | $M_{odt}^{skill}$ (1) | $\frac{M_{odt}^{skill}}{M_{odt}}$ (2) | $M_{odt}^{res}$ (3) | $\frac{M_{odt}^{res}}{M_{odt}}$ (4) | $M_{odt}^{nat}$ (5) |
|---|---|---|---|---|---|
| $MPS_{dt-i}^{skill}$ | -0.110 | 0.0819*** | -0.109*** | -0.0955*** | 0.00190 |
| | (0.0818) | (0.0271) | (0.0278) | (0.00769) | (0.00347) |
| $MPS_{dt-i}^{res}$ | -0.195*** | -0.160*** | -0.104*** | -0.0386 | -0.0168*** |
| | (0.0720) | (0.0473) | (0.0319) | (0.0247) | (0.00639) |
| $MPS_{dt-i}^{nat}$ | 0.737*** | 0.377*** | 0.00255 | 0.0857 | 0.0870*** |
| | (0.225) | (0.107) | (0.0793) | (0.0800) | (0.0192) |
| $MPR_{dt-i}$ | 0.719 | 1.272*** | -0.251 | -0.720** | -0.0110 |
| | (0.458) | (0.353) | (0.276) | (0.362) | (0.0644) |
| $lnGDPpc_{odt-1}$ | 7.842 | 37.91*** | 9.776*** | 1.420 | 2.439*** |
| | (16.54) | (8.081) | (1.988) | (1.302) | (0.605) |
| $lnMigStock_{odt-1}$ | 0.651*** | 0.109 | 0.512*** | 0.0403 | 0.192*** |
| | (0.178) | (0.165) | (0.0915) | (0.107) | (0.0292) |
| $lnUnemp_{dt-1}$ | -1.684 | -0.582 | -0.464 | -0.436 | -0.460*** |
| | (1.146) | (0.389) | (0.438) | (0.409) | (0.0781) |
| $lnInterdec9050_{dt-1}$ | 4.083 | -22.92*** | | | 1.721*** |
| | (14.25) | (7.370) | | | (0.632) |
| Constant | -10.43 | -47.34*** | -7.318** | -0.895 | 3.756*** |
| | (25.15) | (10.09) | (3.249) | (2.534) | (0.744) |
| Origin-time FE | yes | yes | yes | yes | yes |
| Origin-dest FE | yes | yes | yes | yes | yes |
| Observations | 1,020 | 1,020 | 878 | 878 | 27,839 |
| Pseudo $R^2$ | 0.988 | 0.123 | 0.992 | 0.210 | 0.979 |

Notes: Standard errors are clustered at the destination-time level. ***, **, and * indicates significance at the 1%, 5%, and 10% level. The number of lags, $i$, is 5 in columns 1 through 4 and is 1 in column 5.

**Table A-5:** Alternative skill indicator based on **?**

| | $M^{skill}_{odt}$ (1) | $\frac{M^{skill}_{odt}}{M_{odt}}$ (2) | $M^{res}_{odt}$ (3) | $\frac{M^{res}_{odt}}{M_{odt}}$ (4) | $M^{nat}_{odt}$ (5) |
|---|---|---|---|---|---|
| Parsons$_{dt-i}$ | 8.595$^{***}$ | 5.292$^{**}$ | 0.344 | -10.80$^{***}$ | -0.718$^{***}$ |
| | (1.142) | (2.123) | (1.068) | (1.287) | (0.263) |
| $MPS^{res}_{dt-i}$ | -1.226$^{***}$ | -0.203 | 0.229$^{***}$ | 0.251$^{***}$ | 0.0263$^{***}$ |
| | (0.043) | (0.128) | (0.008) | (0.012) | (0.007) |
| $MPS^{nat}_{dt-i}$ | -0.0192 | 0.152$^{*}$ | -0.267$^{***}$ | -0.131$^{***}$ | 0.0465 |
| | (0.037) | (0.089) | (0.058) | (0.022) | (0.031) |
| $MPR_{dt-i}$ | -9.537$^{***}$ | -4.812$^{***}$ | 0.296 | 9.830$^{***}$ | 0.533$^{***}$ |
| | (0.833) | (1.539) | (1.086) | (1.474) | (0.112) |
| $lnGDPpc_{odt-1}$ | 132.7$^{***}$ | 36.94$^{***}$ | | | 1.106 |
| | (4.072) | (10.72) | | | (0.903) |
| $lnMigStock_{odt-1}$ | -0.218 | -0.518$^{***}$ | -0.186 | 0.916$^{***}$ | 0.0570 |
| | (0.235) | (0.107) | (0.192) | (0.147) | (0.0399) |
| $lnUnemp_{dt-1}$ | | | | | -0.251$^{***}$ |
| | | | | | (0.0694) |
| $lnInterdec9050_{dt-1}$ | | | | | 0.955 |
| | | | | | (1.140) |
| Constant | -246.6$^{***}$ | -58.47$^{***}$ | 10.04$^{***}$ | -16.83$^{***}$ | 6.561$^{***}$ |
| | (10.74) | (18.20) | (3.008) | (2.064) | (1.436) |
| Origin-nest-time FE | yes | yes | yes | yes | yes |
| Origin-dest FE | yes | yes | yes | yes | yes |
| Observations | 466 | 466 | 344 | 344 | 9,177 |
| Pseudo $R^2$ | 0.997 | 0.140 | 0.994 | 0.170 | 0.995 |

Notes: Standard errors are clustered at the destination-time level. $^{***}$, $^{**}$, and $^{*}$ indicates significance at the 1%, 5%, and 10% level. The number of lags, $i$, is 5 in columns 1 through 4 and is 1 in column 5.

**Table A-6:** Baseline specification restricted to Parsons sample

| | $M^{skill}_{odt}$ (1) | $\frac{M^{skill}_{odt}}{M_{odt}}$ (2) | $M^{res}_{odt}$ (3) | $\frac{M^{res}_{odt}}{M_{odt}}$ (4) | $M^{nat}_{odt}$ (5) |
|---|---|---|---|---|---|
| $MPS^{skill}_{dt-i}$ | 0.0713*** | 0.0439** | 0.00296 | -0.0929*** | -0.00136 |
| | (0.00948) | (0.0176) | (0.00918) | (0.0111) | (0.00443) |
| $MPS^{res}_{dt-i}$ | -0.931*** | -0.0217 | 0.236*** | 0.0457*** | 0.0225*** |
| | (0.0689) | (0.0892) | (0.0204) | (0.0169) | (0.00861) |
| $MPS^{nat}_{dt-i}$ | -0.0192 | 0.152* | -0.267*** | -0.131*** | 0.0465 |
| | (0.0366) | (0.0892) | (0.0577) | (0.0218) | (0.0338) |
| $MPR_{dt-i}$ | -3.365*** | -1.011*** | 0.527 | 2.588*** | 0.496*** |
| | (0.148) | (0.176) | (0.408) | (0.721) | (0.129) |
| $lnGDPpc_{odt-1}$ | 120.6*** | 29.50*** | | | 0.443 |
| | (5.131) | (8.922) | | | (0.853) |
| $lnMigStock_{odt-1}$ | -0.218 | -0.518*** | -0.186 | 0.916*** | 0.0291 |
| | (0.235) | (0.107) | (0.192) | (0.147) | (0.0411) |
| $lnUnemp_{dt-1}$ | | | | | -0.324*** |
| | | | | | (0.0714) |
| $lnInterdec9050_{dt-1}$ | | | | | 1.503 |
| | | | | | (1.155) |
| Constant | -226.2*** | -48.29*** | 9.815*** | -10.53*** | 7.323*** |
| | (12.54) | (15.90) | (2.331) | (1.389) | (1.616) |
| Origin-nest-time FE | yes | yes | yes | yes | yes |
| Origin-dest FE | yes | yes | yes | yes | yes |
| Observations | 466 | 466 | 344 | 344 | 9,177 |
| Pseudo $R^2$ | 0.997 | 0.140 | 0.994 | 0.170 | 0.995 |

Notes: Standard errors are clustered at the destination-time level. ***, **, and * indicates significance at the 10%, 5%, and 1% level. The number of lags, $i$, is 5 in columns 1 through 4 and is 1 in column 5.

**Table A-7:** Increasing the Lags by 2

| | $M_{odt}^{skill}$ (1) | $\frac{M_{odt}^{skill}}{M_{odt}}$ (2) | $M_{odt}^{res}$ (3) | $\frac{M_{odt}^{res}}{M_{odt}}$ (4) | $M_{odt}^{nat}$ (5) |
|---|---|---|---|---|---|
| $MPS_{dt-i}^{skill}$ | 0.0396 | 0.191*** | -0.0358** | -0.0662*** | 0.000411 |
| | (0.0702) | (0.0299) | (0.0154) | (0.0172) | (0.00436) |
| $MPS_{dt-i}^{res}$ | 0.174*** | 0.0336 | 0.157*** | 0.112*** | -0.0179** |
| | (0.0502) | (0.0371) | (0.0107) | (0.0215) | (0.00761) |
| $MPS_{dt-i}^{nat}$ | 0.422** | 0.186 | -0.378*** | -0.148* | 0.0152 |
| | (0.196) | (0.119) | (0.0551) | (0.0851) | (0.0199) |
| $MPR_{dt-i}$ | 1.457*** | 1.136*** | -0.236** | -0.583*** | 0.0210 |
| | (0.373) | (0.228) | (0.119) | (0.218) | (0.0743) |
| $lnGDPpc_{odt-1}$ | 0.471 | 31.78*** | 2.541*** | 2.409 | 2.776*** |
| | (12.86) | (5.328) | (0.976) | (2.287) | (0.620) |
| $lnMigStock_{odt-1}$ | -0.330* | -0.211* | 0.0452 | -0.000373 | 0.195*** |
| | (0.187) | (0.117) | (0.0408) | (0.123) | (0.0373) |
| $lnUnemp_{dt-1}$ | -1.456 | -0.842** | 0.345 | 0.807* | -0.529*** |
| | (0.898) | (0.336) | (0.223) | (0.426) | (0.0910) |
| $lnInterdec9050_{dt-1}$ | -1.081 | -15.62*** | | | 1.339* |
| | (6.077) | (2.785) | | | (0.794) |
| Constant | 13.42 | -39.57*** | 4.960*** | -5.428 | 3.902*** |
| | (22.00) | (7.947) | (1.735) | (3.991) | (0.989) |
| Origin-nest-time FE | yes | yes | yes | yes | yes |
| Origin-dest FE | yes | yes | yes | yes | yes |
| Observations | 1,006 | 1,006 | 874 | 874 | 27,005 |
| Pseudo $R^2$ | 0.991 | 0.131 | 0.995 | 0.214 | 0.984 |

Notes: Standard errors are clustered at the destination-time level. ***, **, and * indicates significance at the 10%, 5%, and 1% level. The number of lags, $i$, is 7 in columns 1 through 4 and is 3 in column 5.

**Table A-8:** Origin-time clustered standard errors

| | $M_{odt}^{skill}$ (1) | $\frac{M_{odt}^{skill}}{M_{odt}}$ (2) | $M_{odt}^{res}$ (3) | $\frac{M_{odt}^{res}}{M_{odt}}$ (4) | $M_{odt}^{nat}$ (5) |
|---|---|---|---|---|---|
| $MPS_{dt-i}^{skill}$ | -0.00180 | 0.127*** | -0.00531 | -0.0293* | -0.00237 |
| | (0.0418) | (0.00994) | (0.00933) | (0.0164) | (0.00251) |
| $MPS_{dt-i}^{res}$ | -0.0885 | 0.108* | 0.217*** | 0.166*** | -0.0250*** |
| | (0.0980) | (0.0606) | (0.0325) | (0.0483) | (0.00565) |
| $MPS_{dt-i}^{nat}$ | 0.777*** | 0.318*** | -0.172 | 0.0886 | 0.0624** |
| | (0.123) | (0.0801) | (0.122) | (0.0860) | (0.0285) |
| $MPR_{dt-i}$ | 1.324*** | 1.081*** | 0.0262 | -0.582*** | -0.0461 |
| | (0.164) | (0.111) | (0.0963) | (0.196) | (0.0662) |
| $lnGDPpc_{odt-1}$ | 16.95*** | 28.87*** | -0.498 | -1.142 | 2.645*** |
| | (6.558) | (4.098) | (1.014) | (2.193) | (0.513) |
| $lnMigStock_{odt-1}$ | 0.100 | -0.175** | -0.0477 | -0.0543 | 0.196*** |
| | (0.191) | (0.0854) | (0.103) | (0.141) | (0.0361) |
| $lnUnemp_{dt-1}$ | -1.122*** | -0.148 | 0.0401 | 0.595** | -0.481*** |
| | (0.282) | (0.117) | (0.196) | (0.291) | (0.0643) |
| $lnInterdec9050_{dt-1}$ | 0.958 | -14.73*** | | | 1.738*** |
| | (4.633) | (2.176) | | | (0.602) |
| Constant | -22.69* | -37.48*** | 9.455*** | -1.444 | 3.704*** |
| | (11.95) | (6.026) | (2.809) | (3.736) | (0.875) |
| Origin-nest-time FE | yes | yes | yes | yes | yes |
| Origin-dest FE | yes | yes | yes | yes | yes |
| Observations | 1,020 | 1,020 | 878 | 878 | 27,839 |
| Pseudo $R^2$ | 0.991 | 0.131 | 0.994 | 0.215 | 0.984 |

Notes: Standard errors are clustered at the origin-time level. ***, **, and * indicates significance at the 10%, 5%, and 1% level. The number of lags, $i$, is 5 in columns 1 through 4 and is 1 in column 5.

**Table A-9:** Three-way fixed effects (origin-time, origin-destination and destination-time)

| | $M_{odt}^{skill}$ (1) | $\frac{M_{odt}^{skill}}{M_{odt}}$ (2) | $M_{odt}^{res}$ (3) | $\frac{M_{odt}^{res}}{M_{odt}}$ (4) | $M_{odt}^{nat}$ (5) |
|---|---|---|---|---|---|
| $MPS_{odt-i}^{nat}$ | -0.233*** | 0.0504 | -0.149*** | 0.0161 | -0.0133 |
| | (0.0714) | (0.0327) | (0.0575) | (0.0792) | (0.0129) |
| $lnMigStock_{odt-1}$ | 0.0947 | -0.163* | 0.186 | 0.0251 | 0.179*** |
| | (0.107) | (0.0927) | (0.127) | (0.182) | (0.0314) |
| Constant | 9.277*** | 0.247 | 6.807*** | -2.334 | 7.472*** |
| | (1.313) | (0.769) | (1.729) | (1.620) | (0.371) |
| Destination-time FE | yes | yes | yes | yes | yes |
| Origin-time FE | yes | yes | yes | yes | yes |
| Origin-destination FE | yes | yes | yes | yes | yes |
| Observations | 1,554 | 1,554 | 994 | 994 | 44,542 |
| Pseudo $R^2$ | 0.991 | 0.138 | 0.993 | 0.215 | 0.979 |

Notes: Standard errors clustered by origin-time in parentheses. *,** and *** indicate significance at the 10%, 5% or 1% levels respectively. The number of lags, $i$, is 5 in columns 1 through 4 and is 1 in column 5.

**Table A-10:** Three-way fixed effects using the baseline sample

| | $M^{skill}_{odt}$ (1) | $\frac{M^{skill}_{odt}}{M_{odt}}$ (2) | $M^{res}_{odt}$ (3) | $\frac{M^{res}_{odt}}{M_{odt}}$ (4) | $M^{nat}_{odt}$ (5) |
|---|---|---|---|---|---|
| $MPS^{nat}_{odt-i}$ | -0.0334 | 0.0329 | -0.106* | 0.0709 | 0.0756*** |
| | (0.0522) | (0.0651) | (0.0602) | (0.0855) | (0.0177) |
| $lnMigStock_{odt-1}$ | -0.228 | -0.138 | 0.246 | 0.0949 | 0.232*** |
| | (0.192) | (0.101) | (0.151) | (0.186) | (0.0307) |
| Constant | 12.78*** | -0.0720 | 5.983*** | -3.115* | 6.512*** |
| | (2.360) | (0.772) | (2.032) | (1.675) | (0.366) |
| Destination-time FE | yes | yes | yes | yes | yes |
| Origin-time FE | yes | yes | yes | yes | yes |
| Origin-destination FE | yes | yes | yes | yes | yes |
| Observations | 1,020 | 1,020 | 878 | 878 | 27,830 |
| Pseudo $R^2$ | 0.995 | 0.136 | 0.993 | 0.212 | 0.984 |

NOTES: Regressions include origin-time, destination-time and origin-destination fixed effects. Standard errors are clustered at the destination-time level. ***, **, and * indicates significance at the 10%, 5%, and 1% level.

**Table A-11:** Alternative nests structure

| | $M_{odt}^{skill}$ (1) | $\frac{M_{odt}^{skill}}{M_{odt}}$ (2) | $M_{odt}^{res}$ (3) | $\frac{M_{odt}^{res}}{M_{odt}}$ (4) | $M_{odt}^{nat}$ (5) |
|---|---|---|---|---|---|
| $MPS_{dt-i}^{skill}$ | -0.0632 | 0.0855*** | -0.0592** | -0.0162 | -0.00612* |
| | (0.0677) | (0.0222) | (0.0230) | (0.0271) | (0.00334) |
| $MPS_{dt-i}^{res}$ | -0.322* | -0.233 | 0.0947** | 0.175** | -0.0102 |
| | (0.190) | (0.198) | (0.0450) | (0.0760) | (0.00660) |
| $MPS_{dt-i}^{nat}$ | 0.487** | 0.273* | -0.0610 | -0.101 | 0.0103 |
| | (0.206) | (0.161) | (0.0806) | (0.0619) | (0.0203) |
| $MPR_{dt-i}$ | -0.378 | 0.751 | -0.891** | -1.123*** | -0.0741 |
| | (1.004) | (0.471) | (0.372) | (0.343) | (0.0806) |
| $lnGDPpc_{odt-1}$ | 46.53** | 63.66*** | 6.176*** | 3.200 | 2.193*** |
| | (22.40) | (20.65) | (1.241) | (2.358) | (0.610) |
| $lnMigStock_{odt-1}$ | 0.420* | 0.102 | 0.431*** | -0.0736 | 0.174*** |
| | (0.241) | (0.167) | (0.0934) | (0.166) | (0.0385) |
| $lnUnemp_{dt-1}$ | -0.372 | 0.208 | 0.775** | 1.053* | -0.567*** |
| | (1.148) | (0.342) | (0.324) | (0.631) | (0.0959) |
| $lnInterdec9050_{dt-1}$ | -20.79 | -34.86*** | | | 1.169 |
| | (12.65) | (5.974) | | | (0.849) |
| Constant | -63.32* | -82.08*** | -5.637** | -7.621* | 5.233*** |
| | (37.35) | (29.30) | (2.538) | (4.632) | (1.104) |
| Origin-altNest-time FE | yes | yes | yes | yes | yes |
| Origin-dest FE | yes | yes | yes | yes | yes |
| Observations | 954 | 954 | 724 | 724 | 27,849 |
| Pseudo $R^2$ | 0.991 | 0.130 | 0.993 | 0.218 | 0.985 |

Notes: Standard errors are clustered at the destination-time level. ***, **, and * indicates significance at the 1%, 5%, and 10% level. The number of lags, $i$, is 5 in columns 1 through 4 and is 1 in column 5.

**Table A-12:** Regressions using original DIOC data

| | $M^{skill}_{odt}$ (1) | $\frac{M^{skill}_{odt}}{M_{odt}}$ (2) | $M^{res}_{odt}$ (3) | $\frac{M^{res}_{odt}}{M_{odt}}$ (4) |
|---|---|---|---|---|
| $MPS^{skill}_{dt-i}$ | 0.103 | 0.280*** | 0.271*** | 0.374*** |
| | (0.0638) | (0.0519) | (0.0661) | (0.0516) |
| $MPS^{res}_{dt-i}$ | -0.00886 | 0.385* | 0.579*** | 0.637*** |
| | (0.0943) | (0.216) | (0.0790) | (0.0702) |
| $MPS^{nat}_{dt-i}$ | 0.934*** | 0.139 | -0.278 | -0.134 |
| | (0.229) | (0.126) | (0.203) | (0.143) |
| $MPR_{dt-i}$ | 1.689*** | 2.721*** | -0.975*** | -1.092*** |
| | (0.431) | (0.654) | (0.317) | (0.331) |
| $lnGDPpc_{odt-1}$ | 14.51* | 17.71 | | |
| | (8.545) | (18.81) | | |
| $lnMigStock_{odt-1}$ | 0.248 | 0.165 | -0.225 | 0.929** |
| | (0.243) | (0.170) | (0.713) | (0.430) |
| $lnUnemp_{dt-1}$ | -1.776*** | -1.918*** | | |
| | (0.250) | (0.439) | | |
| $lnInterdec9050_{dt-1}$ | 6.068 | -30.19 | | |
| | (9.412) | (18.93) | | |
| Constant | -26.65* | -16.88 | -4.789 | -29.02*** |
| | (13.81) | (25.50) | (13.32) | (4.646) |
| Origin-nest-time FE | yes | yes | yes | yes |
| Origin-dest FE | yes | yes | yes | yes |
| Observations | 714 | 714 | 424 | 424 |
| Pseudo $R^2$ | 0.993 | 0.133 | 0.994 | 0.194 |

Notes: Standard errors are clustered at the destination-time level. ***, **, and * indicates significance at the 1%, 5%, and 10% level. The number of lags, $i$, is 5 in columns 1 through 4 and is 1 in column 5.