Online Appendix for

IS THE MEDITERRANEAN THE NEW RIO GRANDE? US AND EU IMMIGRATION PRESSURES IN THE LONG RUN

Gordon Hanson and Craig McIntosh

A1. Estimation of the forecasting model.

We estimate the forecasting model using 2000-2010 migration flows, as in Table 2, except that now we include all seven of the dyad-level gravity variables simultaneously in a single regression. The variables used in the analysis are: $y_{csd,t}$, the percentage of age-gender group c from sending-country s that has migrated to destination-country d as of year t; $\frac{L_{cs,t}}{L_{cd,t}}$, the relative size of age cohort c for sending-country s and destination-country d at time t; $\frac{w_{cs,t}}{w_{cd,t}}$, relative per capita GDP for age cohort c corresponding to dyad sd at time t; and δ_c, δ_s, and δ_d, fixed effects for the cohort, sending country, and destination country.

The estimating equation is,

$$
\begin{aligned}
y_{csd,t+1} - y_{csd,t} &= \delta_c + \delta_s + \delta_d + \alpha\left[\ln\frac{L_{cs,t}}{L_{cd,t}} - \ln\frac{L_{cs,t-1}}{L_{cd,t-1}}\right] + \beta_1\left[\ln\frac{w_{cs,t}}{w_{cd,t}}\right] + \beta_2\left[\ln\frac{w_{cs,t}}{w_{cd,t}}\right]^2 \\
&+ \sum_i \gamma^i \tau_{sd}^i + \sum_i \rho^i\left(\left[\ln\frac{L_{cs,t}}{L_{cd,t}} - \ln\frac{L_{cs,t-1}}{L_{cd,t-1}}\right] * \tau_{sd}^i\right) + \varepsilon_{csd}
\end{aligned}
$$

where the gravity variables τ_{sd}^i $i=1,\ldots,7$, are log distance between the sending and destination country, indicators for the sending and destination country having a common border, common language, and shared colonial history, and indicators for whether the migration rate of over-age-50 migrants from the sending to the destination country in 2000 was in the top 50 percent, top 20 percent, and top 5 percent across all sending-destination dyads.

Using our labor-supply projections derived from UN data and GDP forecasts derived from IMF data (as described in the text), we predict the dyadic decadal migration flows that will occur for the four 10-year age cohorts (15-24, 25-34, 35-44, and 45-54) over four decades (2010-2020, 2020-2030, 2030-2040, and 2040-2050). The labor-supply values that we feed into the forecasts include a mix of cohorts already born and cohorts yet to be born, as of 2015. The oldest already-born cohort is aged 45-54 in 2010 and thus is used in the forecast only once for the change in migration rate over 2010-2020; the youngest already-born cohort is aged 5-14 in 2010 and which is used to forecast migration flows over 2020-2030, 2030-2040, and 2040-2050; and the youngest yet-to-be born cohort is aged 15-24 in 2040 and used to forecast migration flows over 2040-2050. Of the 7 birth cohorts represented in the analysis, two are born after 2015; of the 16 age cohorts used in the forecasts (2 birth cohorts appear in the analysis once, 2 appear 2 times, 2 appear 3 times and 1 appears 4 times), 3 correspond to individuals born after 2015. See note 13 in the main text for a listing of each birth cohort and how many times it appears in the data.

To construct forecasts of migrant stocks, we add these forecasted decadal migration inflows to any pre-existing migrant stocks and proceed to calculate future stocks of migrants from each origin in each destination in each decade. Where we observe stocks of migrants at the beginning of the period, we add the forecasted flows to these stocks. Where we predict flows for new cohorts with no observed migration, we assume that the initial stock of migrants is zero. These predicted stocks can then be summed across gender, age, origin, or destination to calculate totals of foreign-born migrants aged 15-64 for each decade. With zero predicted future migration flows, or predicted decadal flows after 2010 that are smaller than average decadal flows before 2010, forecasted migrant stocks in destination countries decline over time, as larger cohorts of older migrants age out of the 15-64 age group and smaller cohorts of new migrants are incorporated into the stock. With predicted decadal flows after 2010 that are larger than average decadal flows before 2010, forecasted migrant stocks rise over time.

Appendix Table A1: Regression used to fit the prediction model.

Dependent variable is change in % of cohort migrated within dyad, 2000-2010	Coefficient (Standard Error)
1990-2000 change in age-specific	0.295**
ln(origin/dest) birth cohort ratio	(0.133)
Log (origin/dest) GDP ratio in year	0.00272
cohort turned 15	(0.006)
(Log (origin/dest) GDP ratio in year	0.00265***
cohort turned 15) squared	(0.001)
top 50% 2000 migration rate over 50	0.0111**
	(0.006)
top 50% * birth cohort ratio	-0.0389***
	(0.014)
top 20% 2000 migration rate over 50	0.0111
	(0.012)
top 20% * birth cohort ratio	0.121***
	(0.032)
top 5% 2000 migration rate over 50	0.159*
	(0.096)
top 5% * birth cohort ratio	0.675***
	(0.248)
Log Distance from Origin to Destination	-0.0213***
	(0.005)
Log Distance * birth cohort ratio	-0.0381**
	(0.015)
Contiguous Origin & Destination	0.0972
	(0.079)
Contiguous * birth cohort ratio	0.917**
	(0.366)
Colonial Relationship	-0.0504**
	(0.023)
Colonial * birth cohort ratio	-0.178**
	(0.077)
Common Language	0.0622***
	(0.008)
Common Language & birth cohort ratio	0.0547
	(0.035)
Constant	0.555*
	(0.310)
Observations	18,297
R-squared	0.114

Analysis includes fixed effects for 10-year birth cohorts, origin, destination, and gender. Weighted by birth cohort population size to make results representative for all individuals born in the origin countries. Standard errors clustered at the dyad/gender/cohort level. *** p<0.01, ** p<0.05, * p<0.1.

A2. Varying GDP growth in the forecasts.

We are almost certainly able to forecast population more accurately than GDP. This raises the question of the sensitivity of our forecasts to potential future swings in GDP. For example, imagine that Africa were to enter a high-growth period, how much would the resulting increase in labor demand hold back pressures for emigration that our model otherwise suggests will take place? To answer this question, we feed alternate scenarios for GDP growth into our prediction model, examining outcomes if African GDP grew 10%, 25%, and 50% faster than the IMF forecast currently predicts.

Examination of Table A1 already suggests that the effect may be modest, but positive; meaning that higher income ratios would encourage *higher* emigration. The coefficients on GDP ratios are small, and indicate a convex relationship meaning that the effect will be smallest for countries such as those in SSA who would begin a change in income ratios from a low base.

Indeed, when we examine the migration predictions produced by the different GDP runs, we find a trivially small difference between them. Table A2 shows that while we do get a noticeable difference in prediction values when we move from a model without GDP to one with the linear and quadratic function of the GDP ratios, varying the economic growth rate of countries in Sub-Saharan Africa by wide margins has little effect on migration rates. This is largely because GDP ratios have only a very weakly positive relationship with migration when the income gap between origin and destination is large.

Appendix Table A2: Predicted Migration when Shifting GDP Forecasts.

Correlation between forecasts of migrant stocks for SSA from 2020 on:

	No GDP	**GDP ratio linear & quadratic**	African GDP up by 10%	African GDP up by 25%	African GDP up by 50%
No GDP	1				
GDP ratio linear & quadratic	0.9741	1			
African GDP up by 10%	0.9735	1	1		
African GDP up by 25%	0.9726	0.9999	1	1	
African GDP up by 50%	0.9713	0.9997	0.9998	0.9999	1

Note: 'GDP ratio linear & quadratic' is the specification used in the paper's main tables.