



DEMOGRAPHIC RESEARCH

A peer-reviewed, open-access journal of population sciences

DEMOGRAPHIC RESEARCH

**VOLUME 45, ARTICLE 4, PAGES 87–140
PUBLISHED 8 JULY 2021**

<http://www.demographic-research.org/Volumes/Vol45/4/>

DOI: 10.4054/DemRes.2021.45.4

Research Article

**Global bilateral migration projections
accounting for diasporas, transit and return
flows, and poverty constraints**

Albano Rikani

Jacob Schewe

© 2021 Albano Rikani & Jacob Schewe.

This open-access work is published under the terms of the Creative Commons Attribution 3.0 Germany (CC BY 3.0 DE), which permits use, reproduction, and distribution in any medium, provided the original author(s) and source are given credit.

See <https://creativecommons.org/licenses/by/3.0/de/legalcode>

Contents

1	Introduction	88
2	Methods	90
2.1	Model	90
2.2	Parameter estimation	92
2.2.1	Data	92
2.2.2	Fitting procedure	93
2.3	Simulations	95
2.3.1	Scenarios	95
2.3.2	Natural population change	95
3	Results	96
3.1	Estimation results	96
3.2	Past migration trends	99
3.3	Future migration trends	101
4	Discussion and conclusions	105
5	Acknowledgments	109
	References	110
	Appendix: Global bilateral migration projections accounting for diasporas, transit and return flows, and poverty constraints	115
	A. Model coverage	115
	Table A-1	116
	Figure A-1	117
	Figure A-2	118
	Figure A-3	119
	Figure A-4	120
	Figure A-5	121
	Figure A-6	122
	Figure A-7	123
	Figure A-8	125
	Figure A-9	126
	Figure A-10	127
	Figure A-11	128
	Figure A-12	129
	Figure A-13	130
	Figure A-14	131
	Figure A-15	132
	Figure A-16	133
	Figure A-17	135
	Figure A-18	136
	Figure A-19	137
	Figure A-20	138
	Figure A-21	139

Global bilateral migration projections accounting for diasporas, transit and return flows, and poverty constraints

Albano Rikani¹

Jacob Schewe²

Abstract

BACKGROUND

Anticipating changes in international migration patterns is useful for demographic studies and for designing policies that support the well-being of those involved. Existing forecasting methods do not account for a number of stylized facts that emerge from large-scale migration observations and theories: existing migrant communities – diasporas – act to lower migration costs and thereby provide a mechanism of self-amplification; return migration and transit migration are important components of global migration flows; and poverty constrains emigration.

OBJECTIVE

Here we present hindcasts and future projections of international migration that explicitly account for these nonlinear features.

METHODS

We develop a dynamic model that simulates migration flows by origin, destination, and place of birth. We calibrate the model using recently constructed global datasets of bilateral migration.

RESULTS

We show that the model reproduces past patterns and trends well based only on initial migrant stocks and changes in national incomes. We then project migration flows under future scenarios of global socioeconomic development.

CONCLUSIONS

Different assumptions about income levels and between-country inequality lead to markedly different migration trajectories, with migration flows either converging towards net zero if incomes in presently poor countries catch up with the rest of the world; or remaining high or even rising throughout the 21st century if economic development is

¹ Potsdam Institute for Climate Impact Research and the University of Potsdam, Potsdam, Germany.

² Potsdam Institute for Climate Impact Research, Potsdam, Germany. Email: jacob.schewe@pik-potsdam.de.

slower and more unequal. Importantly, diasporas induce significant inertia and sizable return migration flows.

CONTRIBUTION

Our simulation model provides a versatile tool for assessing the impacts of different socioeconomic futures on international migration, accounting for important nonlinearities in migration drivers and flows.

1. Introduction

The United Nations estimated more than 270 million international migrants globally in 2019 (United Nations 2019a). Migrants thus make up about 3.5% of the world's population. Migration transforms both sending and receiving countries through the transfer of labor force, skills, money, and cultural norms and values (Willekens et al. 2016; Bove and Elia 2017).

Potential future trends in global migration are therefore of interest. A number of different forecasting methods have been developed to predict future migration flows between countries or regions; relying, for instance, on time-series extrapolation (De Beer 1993), expert elicitation (Lutz, Sanderson, and Scherbov 1998), or combinations of both in Bayesian frameworks (Bijak and Wiśniowski 2010; Bijak et al. 2019). Such methods can produce probabilistic forecasts that account for different sources of uncertainty and make extensive use of past migration statistics to inform likely future outcomes (Azose and Raftery 2015; Azose, Ševčíková, and Raftery 2016). However, they typically do not account explicitly for the drivers of migration, and thus do not lend themselves to exploring the impacts of possible future changes in these drivers.

On the other hand, scenario-based projection methods are designed to quantify the influence of different assumptions about drivers and processes on the estimates of future migration. Recent examples of global migration projections include those produced in the context of quantifying the Shared Socioeconomic Pathways (SSP, O'Neill et al. 2017), a set of narratives describing alternative futures that are widely used in global change research. In a set of country-level population projections under the SSPs, international migration was assumed to follow past emigration and immigration rates until mid-century and then converge toward zero net migration (KC and Lutz 2014). Alternative, nonzero net migration scenarios were based on scaling net migration rates by each country's share of global wealth, relative to a baseline scenario where past net migration rates were assumed to persist (Abel 2018). Another study used past gross migration rates between sending and receiving countries or regions, instead of net migration rates, and proposed scenarios based on a harmonic mean adjustment between past emigration and

immigration rates (Buettner and Muenz 2018). Finally, a gravity model for migration was applied in a stylized macroeconomic model to project future migration in dependence of changes in education and gross domestic product (GDP) per capita (Docquier 2018).

This paper aims to contribute to the migration projection literature by developing projections of bilateral migration conditional on existing scenarios for major drivers of migration, and exploiting the empirical relationships between these drivers and migration flows in the past. We present a projection model that accounts for several key features of migration flows which have been absent in previous projections; while resting on just a small number of variables, for which future scenarios are readily available.

In particular, natural population change, average incomes in origin and destination countries, and the size of migrant communities in the destination (diasporas), have been identified as major determinants of long-term levels and trends in global migration flows (Massey et al. 1993; Hatton and Williamson 2005; Beine 2016). Natural population change affects the size and composition of a population, and thereby affects the number of potential migrants in an origin country, or the demand for labor migrants in a destination country. Incomes in the destination represent the economic gains a migrant may expect from migration. Changing income levels in the origin are thought to have multiple effects. One, put simply, is that people with less income are more inclined to engage in labor migration in order to improve their livelihoods, or those of their families. Another is that very low incomes may prevent people from migrating internationally even if they would like to, for a lack of resources to cover the costs of migration. These two counteracting effects imply a nonlinear, inverse U-shaped relation between emigration and incomes (or development more generally), sometimes called “migration hump,” where the highest emigration rates are observed from middle-income countries (Hatton and Williamson 2005; de Haas 2010, 2007; Clemens 2014; Williamson 2015: see also Appendix, Figure A-1). Finally, migrant communities play an important role by providing information and support to new arrivals (Beine, Docquier, and Özden 2011; Beine 2016; Migali et al. 2018). The diaspora size can thus provide a positive feedback on migration flows: Immigration during one period increases the bilateral migrant stock, which in turn eases, or even stimulates (e.g., through family reunification), further immigration during the next period. The balance between the rate of immigration and the rate of assimilation of migrants into the host society determines the strength of this feedback (Collier 2013).

Few existing projections make use of these empirically established roles of major migration drivers (an exception being Docquier 2018). None, to our knowledge, account for the feedback induced by diasporas, or for poverty constraints and the “migration hump” induced by the complex effects of origin incomes. Neither do existing projections consider transit and return migration flows. These flows can be large (Azose and Raftery 2019), in which case net migration is a poor indicator of actual bilateral flows, and a model neglecting return flows is prone to overestimating migrant stocks.

In this paper, we present a dynamic model of global bilateral migration that accounts for these mechanisms. By simulating flows separately for each combination of origin, destination, and place of birth, individual migrant communities are explicitly represented and changes in their size feed back on flows. The concept of the diaspora effect is expanded to also account for transit migration flows. Emigration from poor countries is constrained in accordance with observed emigration rates. Return migration is modeled explicitly as a function of migrant stocks. We calibrate the model on a global dataset of bilateral flows; demonstrate its performance for past levels and trends in migration; and project future migration flows under five different Shared Socioeconomic Pathways.

2. Methods

2.1 Model

We model the number of migrants of place of birth k moving from country i to country j ($i \neq j, j \neq k$) in a given period as:

$$M_{k,i \rightarrow j} = a \cdot F(G_i) \cdot g_j^{\alpha_g} p_{k,j}^{\alpha_p} d_{ij}^{\alpha_d} P_{k,i}, \quad (1)$$

where $P_{k,i}$ is the number of people born in k and residing in i at the start of the period, and $d_{ij} = d_{ji}$ is the geographical distance between i and j . $p_{k,j} = P_{k,j}/P_k$ is the share of $P_{k,j}$ in the total population born in k , P_k . G_i is gross domestic product (GDP) per capita in i , and $g_i = G_i/G_{glob}$ is the same expressed relative to the global average, G_{glob} . a is a constant scaling factor.

The model thus assumes that the migration count is proportional to the number of people available to migrate; and the migration rate, $m_{k,i \rightarrow j} = M_{k,i \rightarrow j}/P_{k,i}$, depends (potentially nonlinearly) on the distance as well as on relative measures of the diaspora and average income at the destination. The diaspora term $p_{k,j}$ implies (if $\alpha_p > 0$) that emigration from i increases with the share of the country's native population already living abroad; and that those countries that host a higher share of the total diaspora will be more popular destinations. Similar relationships are implied in existing empirical models, although absolute measures of diaspora size and income are often used (Beine 2016). Importantly, in equation 1, the size of a given diaspora community $P_{k,j}$ only affects flows of migrants of the same place of birth k ; there is a separate migration "channel" for each origin-destination-place of birth set.

The observed, complex dependence of migration on origin incomes is modeled through the superposition of two terms: A hyperbolic term with larger values at low incomes, describing the effect of origin income on the intent to migrate; and a sigmoidal

term with larger values at high incomes, describing the effect on resources available to finance migration:

$$F(G_i) = F_{intent}(G_i) \cdot F_{resource}(G_i) = \frac{1}{1 + \frac{G_i}{\tilde{G}}} \cdot \frac{1}{1 + e^{-\gamma(G_i - \tilde{G})}}. \quad (2)$$

These two terms result in a hump-shaped dependence of emigration rates on origin GDP per capita in terms of the parameters \tilde{G} and G (Figure 1, inset).

For $k = i$, equation 1 represents emigration of people from their country of birth (CoB). Accounting for return migration and transit migration, too, is important not only for internal consistency of the model, but also because these flows can be large, e.g., for migrants born in Mexico returning from the United States (Azose and Raftery 2019), and because they affect diaspora size and thus, in turn, future emigration from the CoB. We hypothesize that transit migration ($i \neq k \neq j$) can be described by the same model as emigration from the CoB. A migrant of place of birth k residing in country i may choose to move to a third country j motivated by low attainable incomes in i , a high expected income in j , and/or the presence of a sizeable diaspora in j ; and may be constrained by insufficient income in i to finance migration. We note though that this ignores the fact that migrants often tend to have lower incomes than the native population, and thus may be more inclined to migrate and/or more resource-constrained.

While emigration from CoB and transit migration are thus described by equation 1, we assume that return migration simply depends on diaspora size and distance:

$$M_{j,i \rightarrow j} = b \cdot d_{ij}^{\beta_d} P_{j,i}. \quad (3)$$

This assumption, while simplistic, is in line with a strong proportionality between return flow and diaspora size in empirical estimates of return migration (Azose and Raftery 2019; Abel and Cohen 2019) (Appendix, Figure A-2), as well as with previous findings suggesting that economic factors have limited influence on return migration flows (Constant 2020; Battistella 2018).

A dynamic simulation of migration over time requires that population stocks are simultaneously updated. Apart from migration, population stocks are subject to natural change due to births and deaths, which we account for through country-specific fertility and mortality rates, r_i^* and r_i^\dagger , respectively. Thus, a given population $P_{k,i}$ evolves as:

$$P_{k,i}(t+1) = \begin{cases} \tilde{P}_{k,i}(t) \cdot (1 - r_k^\dagger), & k \neq i \\ \tilde{P}_{i,i}(t) \cdot (1 + r_i^* - r_i^\dagger) + \sum_{l \neq i} \tilde{P}_{l,i}(t) \cdot r_l^*, & k = i, \end{cases} \quad (4)$$

where $\tilde{P}_{k,i}(t) = P_{k,i}(t) - \sum_{l \neq i} M_{k,i \rightarrow l} + \sum_{l \neq i} M_{k,l \rightarrow i}$ is the population size at the end of the previous period plus net migration during the current period. That is, children born to immigrants are added to the native population, in line with the definition of migrants by country of birth adopted here. The natural population change rates are expressed per model time step.

2.2 Parameter estimation

2.2.1 Data

We use UN Department of Economic and Social Affairs data on historical bilateral migrant stocks (United Nations 2019a) and national total population (United Nations 2019b). Historical country-level GDP comes from the Penn World Tables (PWT) release 8.1 (Feenstra, Inklaar, and Timmer 2015), which are reported in terms of 2005 purchasing power parity (PPP), and are thus consistent with the SSP data that we will use for future projections. Some missing country data are taken from PWT 9.0 after rescaling from 2011 to 2005 PPP (Geiger 2018). Geographical distance between countries is taken from CEPII (Mayer and Zignago 2011). Migration flow data come from a global matrix of bilateral international migrant flows (Abel and Cohen 2019), derived from UN migrant stock data (United Nations 2017) using a pseudo-Bayesian method (Azose and Raftery 2019), which allows simultaneous nonzero flows in both directions of a given migration channel, thereby accounting for return migration flows. We refer to this flow dataset as A19 hereafter. It contains 5-year flows for five periods between 1990 and 2015. Such flow estimates do not represent direct observations, but are the only globally consistent data source available. Comparison with flow data reported by a number of industrialized countries (United Nations 2015) shows that inconsistencies in those data – e.g., between reports by the sending and receiving country for a single flow – sometimes are of similar size as the differences between the reported flows and A19. While there have been attempts to harmonize the reported flow data e.g., by using multivariate models and expert knowledge, these synthetic datasets are only available for flows to and from relatively small sets of highly developed countries (Raymer et al. 2013).

The UN migrant stock data generally include refugees, either because countries include refugees in their population censuses or because refugee populations have been added to the total migrant stocks for countries that are deemed to not report refugee populations as part of their migrant statistics (United Nations 2017). Therefore, the A19 flow data also generally include flows of refugees. In contrast, our equation 2, which describes emigration as a function of origin GDP, is not thought to be a good representation of refugee flows. For estimating the parameters of equation 2 we therefore make an attempt to remove refugee flows from the A19 total flows in cases where separate data on bilateral refugee stocks are available from the UN High Commissioner for Refugees

(UNHCR 2020). Since the UNHCR reports bilateral refugee stocks by country of origin – rather than by country of birth or citizenship – we can only account for direct flows between origin and destination country, ignoring refugees’ (unknown) country of birth. The refugee flows are derived as the difference in refugee stocks between two adjacent periods, accounting for natural population change and naturalization of refugees:

$$\hat{M}_{i \rightarrow j}^{\text{rfg}}(t) = P_{i,j}^{\text{rfg}}(t+1) - P_{i,j}^{\text{rfg}}(t) \cdot (1 + r_i^* - r_i^\dagger) + N_{i,j}^{\text{rfg}}(t), \quad (5)$$

where $P_{i,j}^{\text{rfg}}$ is the refugee population in j with origin country i , and $N_{i,j}^{\text{rfg}}$ the number of refugees that became naturalized during the 5-year period (both from UNHCR 2020). Negative values of $M_{i,i \rightarrow j}^{\text{rfg}}(t)$ are assumed to be return flows, and thus figure into the flow from j to i . Thus, we modify the A19 bilateral flow estimates $\hat{M}_{i \rightarrow j}$ as follows to approximate nonrefugee flows:

$$\hat{M}_{i \rightarrow j}^{\text{norfg}} = \hat{M}_{i \rightarrow j} - \max(0, \hat{M}_{i \rightarrow j}^{\text{rfg}}) - \min(0, \hat{M}_{j \rightarrow i}^{\text{rfg}}). \quad (6)$$

Equation 2 is also not assumed to hold for (nonrefugee) return migration, which is instead described by equation 3. For the estimation of equation 2, we therefore exclude, from the A19 dataset, flows where the origin country is more than twice as rich as the destination country ($G_{\text{origin}} > 2 \cdot G_{\text{destination}}$); assuming that such flows represent predominantly return migration. While this simple criterion is by no means precise, it does remove from the dataset some of the most important return migration routes, such as from the United States to Mexico or from the Gulf States to South Asia, while preserving the bulk of migration flows in the dataset (Appendix, Table A-1 and Figure A-3). The exact choice of the threshold has little influence on the estimated parameters (Appendix, Figure A-4).

2.2.2 Fitting procedure

Let $\hat{M}_{i \rightarrow j}$ be the matrix of bilateral flows reported in A19 (note that these reported flows are not broken down by country of birth). We proceed in two steps. Using Nonlinear Least Squares, we first estimate the parameters of equation 2, the hump-shaped relationship between emigration and origin GDP per capita, from gross emigration rates. To this end, after removing refugee and return flows as described above, we aggregate the flow data over all destination countries, to obtain total gross emigration flows by origin country and period.

We fit equation 2 to the resulting total emigration flows:

$$\sum_{l \neq i} \hat{M}_{i \rightarrow l}^{\text{norfg}} \approx a_e \cdot F(G_i) = a_e \cdot \frac{1}{1 + \frac{G_i}{\bar{G}}} \frac{1}{1 + e^{-\gamma(G_i - \bar{G})}}. \quad (7)$$

In a second step, we fit Equation 1 and 3 to the full dataset of bilateral flows from A19, i.e., without removing any refugee or return flows. To simplify computation, we make the approximating assumption that the observed flow $\hat{M}_{i \rightarrow j}$ is composed only of emigration and return migration. This neglects transit flows, which constitute 9% of global flows (Azose and Raftery 2019), but is still more detailed than previous empirical studies that typically model exclusively emigration flows from sending countries (Beine 2016; Wesselbaum and Aburn 2019).

$$\hat{M}_{i \rightarrow j} \approx M_{i, i \rightarrow j} + M_{j, i \rightarrow j} \approx a \cdot F(G_i) \cdot g_j^{\alpha_g} p_{i,j}^{\alpha_p} d_{ij}^{\alpha_d} P_{i,i} + b \cdot d_{ij}^{\beta_d} P_{j,i}, \quad (8)$$

where we use the parameter values estimated in the first step for $F(G_i)$.

Robustness test. To obtain alternative estimates for the parameters of Equation 1 and 3, we derive a second bilateral flow dataset, using a simple stock-differencing approach, rather than the pseudo-Bayesian demographic accounting approach (Abel and Cohen 2019). Taking migrant stock data from the latest release of the UN database (United Nations 2019a), we employ the method described as “stock differencing, reverse negative” in (Abel and Cohen 2019), but additionally account for natural changes due to deaths (births do not count towards foreign migrant stocks since newborns have the host country’s place of birth): $\Delta P_{i,j}(t) = P_{i,j}(t+1) - P_{i,j}(t) \cdot (1 - r_i^\dagger)$, where r_i^\dagger is the mortality rate expressed per 5-year time period. Again neglecting transit flows, we assume that increases in bilateral stocks are due to immigration and decreases in stocks are due to return migration, and fit the following equation to the derived bilateral flows.

$$\Delta P_{i,j} = \begin{cases} a \cdot F(G_i) \cdot g_j^{\alpha_g} p_{i,j}^{\alpha_p} d_{ij}^{\alpha_d} P_{i,i}, & \text{if } \Delta P_{i,j} > 0 \\ -b \cdot d_{ij}^{\beta_d} P_{i,j}, & \text{if } \Delta P_{i,j} < 0. \end{cases} \quad (9)$$

We note that the stock-differencing approach is prone to underestimating real bilateral flows, since observed changes in stocks may be caused by a combination of immigration and return migration flows, rather than only one or the other. Moreover, the bilateral flow matrix estimated with this approach is incomplete, because many bilateral stocks are not reported individually in the UN data but aggregated into “Other North” and

“Other South” categories. Thus, the estimation from this second flow dataset serves as a robustness test only.

Choice of estimator: Many empirical migration studies estimate gravity models in logarithmic form. Because the traditional gravity model is log-linear, this then allows for using simple linear regression analysis. Given the relatively complex form of our model, compared to common gravity-type empirical models, log-transformation does not simplify the estimation problem here. We therefore use nonlinear Least Squares to estimate the untransformed model equations. This also circumvents problems arising from zero-flow values in both dependent (flows) and independent variables (diaspora), which might need to be excluded from the analysis, and therefore could lead to biased estimates, in a log-transformed model (Santos Silva and Tenreyro 2006).

2.3 Simulations

We simulate³ directed migration flows – defined here as five-year transitions, in line with the historical data used for calibration – by origin, place of birth, and destination. Our dynamic simulation model covers 177 countries or territories; 44 mostly small countries or territories are excluded, mainly because of missing GDP data either for the past or for the SSPs (see Appendix).

2.3.1 Scenarios

Future GDP data come from one of five long-term macroeconomic projections produced by the OECD (Dellink et al. 2017). They are based on the five SSP narratives, which provide qualitative descriptions of possible future societal, economic, political, technological, environmental, and lifestyle changes, among other aspects (O’Neill et al. 2017; Riahi et al. 2017). For the macroeconomic projections, each of these narratives was translated into a set of assumptions about key drivers of economic growth, such as trade openness or education levels (KC and Lutz 2014). The narratives represent a set of alternative, plausible worlds; for instance, a world that is largely oriented towards sustainability, global equality, and human well-being (SSP 1), or one which is more characterized by regional conflicts, low levels of global coordination and cooperation, and declining investments in human capital and innovation (SSP 3; a summary of the main assumptions pertaining to all five narratives is provided in Tables 1–3 in O’Neill et al. (2017)). Consequently, national economies and per-capita incomes develop differently under each set of assumptions. The SSP projections thus allow us to explore, and contrast, the potential effects on

³By simulation, we mean a numerical experiment running our model forward in time; not to be confused with other, specific uses of the term e.g., in Monte Carlo or microsimulations.

global migration of several markedly different, but plausible, future development trajectories.

2.3.2 Natural population change

To account for natural population change, we apply estimated historical natural population fertility and mortality rates from the United Nations (UN) World Population Prospects 2019 (WPP19, United Nations 2019b) to our simulations starting in 1990; and projected natural population fertility and mortality rates from the WPP19's zero-migration projection variant to our simulations starting in 2015. Diasporas are assumed to have the same fertility and mortality rates as their country of birth. We note that this assumption is a strong simplification of real fertility patterns among migrants. In reality, diasporas can sometimes have higher natural change rates than both the origin and destination country populations due to age-specific self-selection. On the other hand, diaspora fertility may also be lower than origin country fertility due to skill-specific self-selection (e.g., Wolf and Mulder 2019). A dataset including migration from and to both high-income and developing countries suggests that for around 40% of sending countries, emigrants have higher fertility rates on average than in their home country; while the opposite is true for the remaining countries (Beine, Docquier, and Schiff 2013). These complexities are not captured in our simple global model. Nonetheless, our assumption that diaspora natural change rates equal those in the origin countries does account for one important empirical observation: the fertility of immigrant populations differs strongly by country of birth (e.g., Dubuc 2012; Milewski 2010; Schmid and Kohls 2010). In terms of modeling results, our assumption also leads to total historical population changes that are more in line with observations than when making the alternative assumption of diasporas growing at the rate of their residence countries (Appendix, Figure A-5).

3. Results

3.1 Estimation results

Estimation results are shown in Table 1. Estimated against the A19 bilateral flow dataset, the model achieves an R^2 of 0.73; notably, without accounting for any unobserved idiosyncrasies of individual migration channels, which are commonly discounted in empirical models through fixed effects. The estimated parameter values are largely consistent with our expectations. Bilateral migration substantially increases with diaspora size and destination GDP. The parameters related to origin GDP are such that expected emigration rates are highest around roughly 5000 US Dollars of GDP per capita (Figure 1), as found previously using a nonparametric regression approach (Clemens 2014). Indeed, the fit

pora effect. Destination GDP has a larger effect than in the main estimation, and return flows depend more strongly on distance. This may be because the incomplete sample of bilateral stocks that the alternative flow dataset is based on is biased towards middle- and high-income destination countries (Appendix, Figure A-6).

Table 1: Parameter values estimated from the A19 bilateral flow data, with (main) and without using the distance variable (no distance); and estimated from the alternative dataset derived through simple stock differencing (stock diff.). Values used for the dynamic simulations are shown in bold

Variable		estimate (main)	estimate (no distance)	estimate (stock difference)
Emigration from CoB and transit migration				
Intercept	a	0.153±0.009	0.270±0.004	0.26±0.04
Diaspora	α_p	0.956±0.003	0.946±0.003	0.952±0.008
Destination GDP	α_g	0.226±0.008	0.252±0.007	0.70±0.04
Distance	α_d	0.079±0.008	0	-0.10±0.02
Original GDP	γ	-0.0015±0.0009		
	\widehat{G}	k\$ 22.6±17.2		
	\widetilde{G}	k\$ 1.2±0.4		
Return migration				
Intercept	b	0.20±0.02	0.127±0.001	153±51
Distance	β_d	-0.06±0.01	0	-1.02±0.05
	N	166530	166530	40462 (emigration) 8361 (return)
	R^2	0.73	0.73	0.46 (emigration) 0.36 (return)

Note: Error margins refer to the 99% confidence interval. Origin GDP parameters are separately estimated from gross emigration flows, as described in the main text; GDP values are real GDP per capita in thousand US dollars at 2005 PPP.

Given the small and inconsistent effects of distance, we choose to simplify our model by setting α_d and β_d to zero, effectively removing the distance terms in equation 1 and 3. Fitting the simplified model again to the A19 dataset yields an equally high R^2 and similar parameter values as the model including distance (Table 1, middle column). We use these parameters to perform dynamic simulations – hindcasts and conditional forecasts – which we discuss in the following. For these dynamic simulations, we iteratively calculate migration flows (using equations 1–3) and stocks (using equation 4) in five-year increments, the results from one period providing input to the next.

In order to reflect parameter uncertainty in the simulation results, we perform additional simulations using the upper and lower bounds of the 66% confidence interval estimated for \widehat{G} , the parameter characterizing the decline of migration intent with rising origin income. Systematic variation of all model parameters shows that uncertainty in

this parameter has by far the largest effect on simulated migration counts, both during the historical period and in the SSP scenarios (Appendix, Figures A-7–A-8).

3.2 Past migration trends

We initialize the model with year-1990 bilateral migrant stocks (United Nations 2019a), and run it until 2015 using historical GDP (Feenstra, Inklaar, and Timmer 2015) and fertility and mortality rates (United Nations 2019b) as inputs. After 25 years, simulated and reported total migrant stocks are still highly correlated (Figure 2).

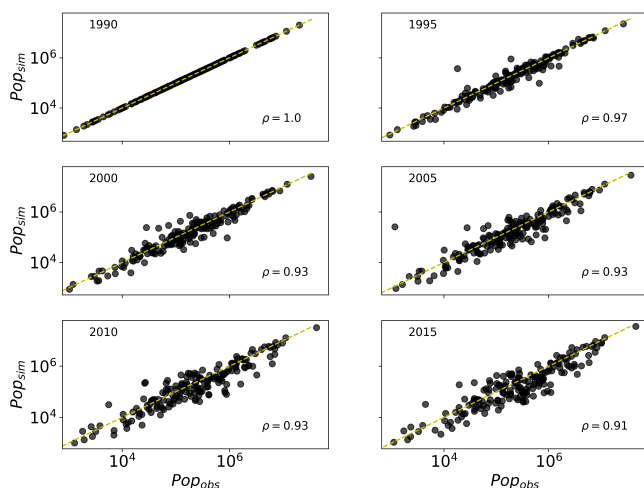
Globally, the model simulates an increase in migration, at a similar rate as observed in the A19 data (Figure 3, black and green lines). While the level of gross migration simulated is somewhat lower than observed, this is at least partly explained by refugee flows, which are included in the observations but not accounted for in our model (note that our method to exclude refugee flows from the A19 data likely does not capture all refugee movements, and thus the green solid line in Figure 3 likely still overstates the size of nonrefugee movements). A counterfactual simulation, with GDP held fixed at 1990 levels (black dotted lines in Figure 3), indicates that the increase in global migration would have been virtually the same even if it had not been for changes in GDP. This is not a trivial finding given the complex effects of origin and destination country GDP on migration; it suggests that while the recent reduction in global economic inequality – the between-country Gini coefficient dropped from 0.75 to 0.65, approximately, between 1990 and 2010 (Dellink et al. 2017) – may have acted to reduce migration, this has been balanced by the increasing effect on emigration of rising incomes in developing countries.

We next consider net migration at regional and country levels. Net migration levels are approximately matched for most world regions (Figure 4; also shown in the Appendix, Figure A-9, in percentage terms relative to the region’s total population) and large immigration and emigration countries (Figures 5–6, Appendix Figures A-10–A-11). We point out again that our model is kept deliberately simple and does not account for any country-specific or bilateral factors other than GDP and bilateral migrant stocks.

The simulation also captures some of the long-term trends in net migration between 1990 and 2015: Rising net emigration in Africa as well as in East, Southeast, and South Asia; and rising net immigration in West Asia, Europe, and Oceania. The counterfactual simulation with GDP held fixed at 1990 levels shows monotonic increases in the magnitude of net migration in almost all regions and countries (even in percentage terms relative to the total population size, Appendix Figures A-9–A-11), illustrating the positive feedback induced by diasporas. Compared to this counterfactual baseline, changes in GDP since 1990 have resulted, especially after 2000, in lower net immigration in Oceania, and lower net emigration from East Asia and Latin America, according to the model; while

they resulted in larger net immigration to West Asia and Europe, and larger net emigration from South Asia and Southeast Asia.

Figure 2: Simulated versus observed migrant stocks per residence country, with Pearson correlation coefficient



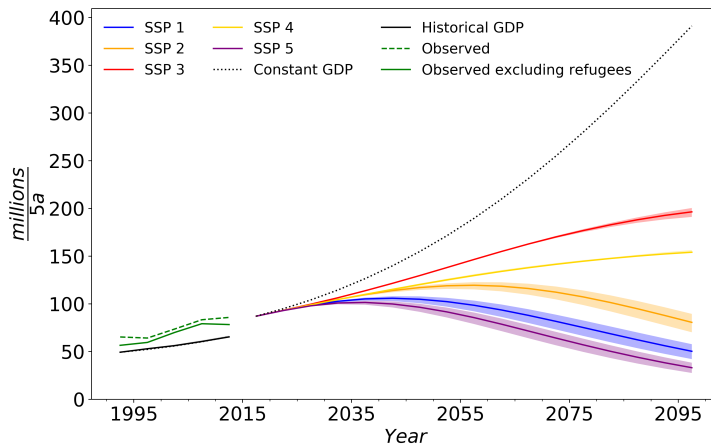
Note: Dashed line shows identity.

The model does not reproduce many of the short-term variations in net migration. This is expected given that the model accounts neither for refugee movements, nor for other sources of short-term variations in migration flows, such as immigration policies or geopolitical changes. For instance, following the collapse of the Soviet Union, large flows are observed between former constituent countries and other countries such as Germany and Israel in the years after 1990, but are not captured by the model. The data also show the massive emigration from Syria to Turkey and other countries since 2011. Finally, estimated net flows from many large Asian emigration countries, to destination countries such as the United Kingdom, Australia, and the Gulf States, spike in the period 2005–2010, before declining again. The simulations do not capture these short-term spikes.

Besides the aggregate measure of net migration, many of the largest individual bilateral flows are also reproduced well (Figure 7). When the simulations are broken down into emigration from CoB, return, and transit flows, they highlight the important role of return and transit migration. While emigration from CoB is the largest component of total migration for most of the large flows, return migration is an important secondary component in many of them (such as flows between Russia and Ukraine, or between India

and Pakistan), and is the largest component in some of them (e.g., the United States to Mexico, Hong Kong to mainland China, India to Bangladesh). Noticeable transit flows are simulated between Russia and Ukraine, as well as from India to Saudi Arabia. High shares of transit migration in total gross migration are simulated in some countries in the Middle East, the Balkans, the Baltics, the Caribbean, and Central America, among others (Figure 8 and Appendix, Figure A-12).

Figure 3: Global gross migration counts



Note: Green dashed line represents A19 data; green solid line is the same but excluding refugee flows. Other lines show two sets of model simulations: One where the model is initialized in 1990 and GDP is either held constant at 1990 levels (black dotted line) or varied according to historical records (black solid line); and another one where the model is initialized in 2015 and GDP is either held constant at 2015 levels (black dotted line) or varied according to the SSPs (Dellink et al. 2017) (colored solid lines). Shading around the simulations with varying GDP represents the 66% confidence interval of the most uncertain model parameter, \hat{G} . Horizontal axis indicates the middle year of each 5-year period.

3.3 Future migration trends

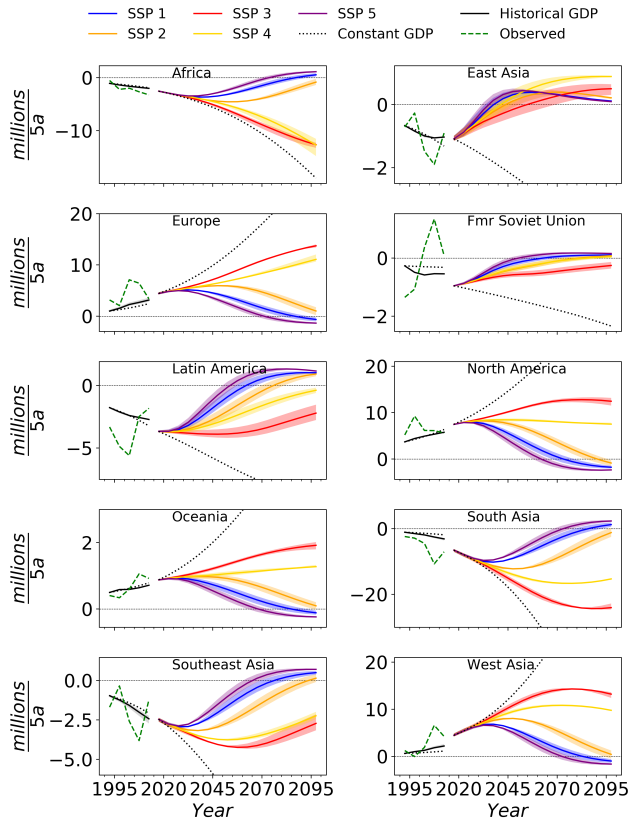
We now initialize the model with year-2015 stocks to generate future scenarios, or conditional predictions, of bilateral migration flows. Because the model is not perfect and does not account, for instance, for refugee movements, the year-2015 migrant stocks resulting from our 25 years of historical simulation are different than observed. Therefore, our future simulations of flows generally start from a different level than where the historical simulation ended. The smaller the discontinuity between the two sets of simulations for a given region or country, the higher the agreement between the historical simulation and

the actual historical evolution of those bilateral migrant stocks that are most relevant for predicting this region's or country's flows.

We run the model until 2100 using projected GDP changes under the five Shared Socioeconomic Pathways (Dellink et al. 2017) and projected natural population change from the UN World Population Prospects 2019 (United Nations 2019b) zero-migration variant. Again, an additional simulation with constant year-2015 GDP serves for comparison. This constant-GDP simulation shows continuously rising net migration throughout the 21st century in all world regions and most large countries (Figures 4–6). The SSP simulations diverge from this. At the level of world regions (Figure 4), all SSPs lead to lower (absolute) net migration flows than the constant-GDP simulation. The projections differ markedly between the different SSPs: Under SSP 1, SSP 5 and, somewhat more slowly, SSP 2, net migration flows approach zero by the end of the century in all world regions and many countries. On the other hand, under SSPs 3 and 4, net migration flows keep rising throughout the century in Africa, Europe, and Oceania. In Southeast Asia, West Asia and, less pronounced, Latin America, SSP 3 leads to a peak around 2060, followed by a decline in net migration flows; the peak is somewhat later in North America and South Asia. In East Asia, SSP 3 and SSP 4 lead to substantial net *immigration* by the end of the century. This is because of declining net emigration from China and relatively high net immigration to Japan and South Korea. Only the Former Soviet Union region exhibits declining net migration under all SSPs.

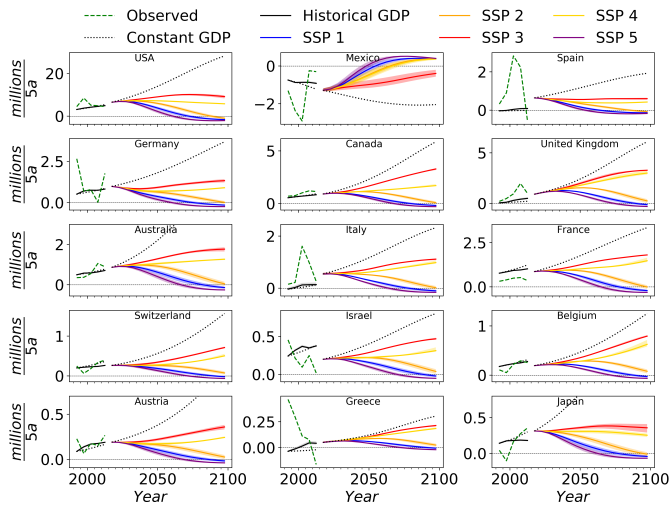
These patterns are mirrored at the country level (Figures 5 and 6). Interesting cases include the South Asian countries Bangladesh and Myanmar, where all the SSPs imply substantially higher net emigration during the coming decades than when assuming constant GDP. The low GDP per capita in these countries today means that a rise in GDP per capita initially leads to strong increases in emigration, according to equation 2; implying that poverty today is a major constraint on emigration from these countries. Over the course of the century, under SSPs 1, 2, and 5, net migration in these countries however peaks and declines again, tracing the “migration hump” described in earlier theoretical and empirical literature (Clemens 2014; European Commission Joint Research Centre and JRC 2018). On the other hand, the much slower economic development in this region in SSP 4 and, particularly, SSP 3 means that a peak is reached only very late, or not at all within the 21st century. Net immigration to Malaysia and South Africa show a similar pattern, reflecting inflows from presently poor South Asian and sub-Saharan African countries, respectively.

Figure 4: Net migration counts in 10 world regions

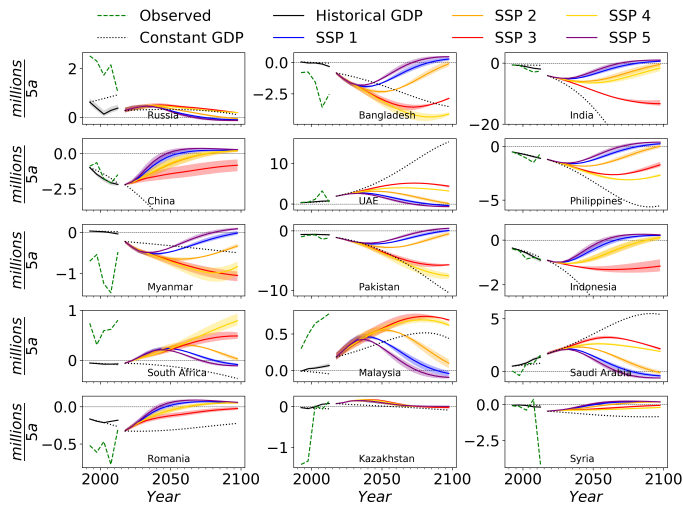


Note: Lines as in Figure 3. Positive net migration means more immigration than emigration.

Figure 5: Net migration counts in OECD countries



Note: As Figure 4 but for individual OECD countries. Countries are sorted in descending order by their average historical net migration count, excluding refugees. Results for net migration rates are shown in the Appendix, also including additional countries (Figures A-5, A-6, and A-11–A-16).

Figure 6: Net migration counts in non-OECD countries

Note: As Figure 5 but for individual non-OECD countries.

Interestingly, as simulated net migration flows approach zero under some of the SSPs, they change their sign; countries that start with net immigration are projected to transition to net emigration, and vice versa. This is because, for a given pair of countries, return migration flows lag emigration flows from the CoB, and become larger than the latter when emigration flows have peaked but migrant stocks are still large (Appendix, Figures A-13 and A-14).

While uncertainty related to the model parameters did not show up strongly in our historical simulations, it has a more substantial effect on the future simulations (Figures 3–6, shading). Especially under SSP 3, for individual regions and countries, the spread between simulations assuming different parameter values can be of the same order as the simulated migration count itself; e.g., for Mexico or China. Nonetheless, scenario-related uncertainty, as encapsulated in the spread between the different SSPs, still dominates the total spread of our projections throughout the 21st century.

4. Discussion and conclusions

SSP 3 represents a world characterized by low levels of GDP per capita and high between-country inequality (Dellink et al. 2017). Conversely, SSPs 1 and 5 represent high aver-

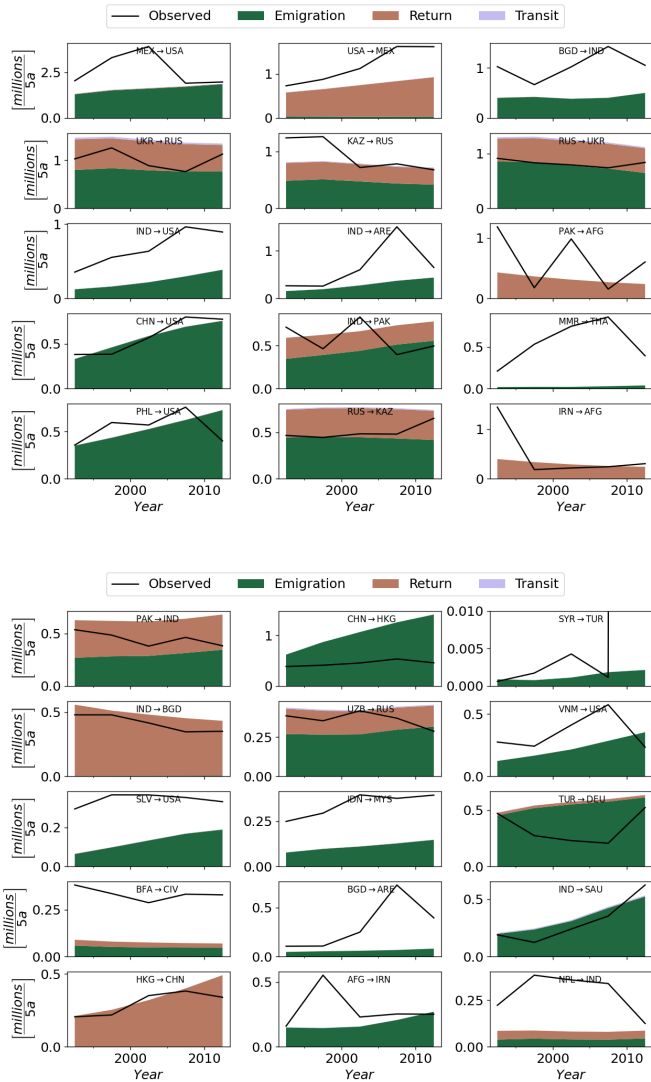
age incomes, stronger economic convergence, and reduced inequality between countries. While global average GDP per capita is similar between SSP 2 and SSP 4 throughout much of the century, it is much more unequally distributed between countries in SSP 4. Our projections highlight the importance of both income levels and between-country income inequality for international migration. The economically stagnant and divided world of SSP 3 implies large and sustained, or even rising, net migration flows from the Global South to the Global North and the Gulf States. Conversely, in the rich and relatively equal worlds of SSP 1 and 5, poor countries are lifted out of poverty constraints, leading to rising migration flows for a few decades; but by 2040, net migration flows are in decline the world over, and by the end of the century, flows are much smaller than today. The divergence between SSP 2 and SSP 4 is strong in many regions, especially in Africa, showing the effects not only of average income levels but also of between-country inequality.

In a given region, our model produces, for the set of five SSPs, a wide range of migration outcomes. However, the range is still narrower than in another set of SSP migration scenarios based on linear scaling with countries' shares in global GDP (Abel 2018). For instance, net migration in Africa does not turn substantially positive even under SSPs 1 and 5 in our simulations. This highlights the relevance of accounting for the major nonlinear driving forces of migration, particularly the reinforcing effect of diasporas, in projecting future migration flows.

Our model also accounts for two counteracting effects of rising origin-country incomes. Consequently, we see initially poor countries tracing the “migration hump” as incomes gradually rise.⁴ Under the more optimistic economic scenarios of SSPs 1 and 5, most countries converge towards near zero net migration by the end of the century or earlier. Net zero migration in our model emerges from an equilibrium between emigration from CoB, transit migration, return migration, and natural population change. It thus does not imply a static world where no one moves, but a world in which emigration flows are generally small enough to be balanced by return migration and natural population growth.

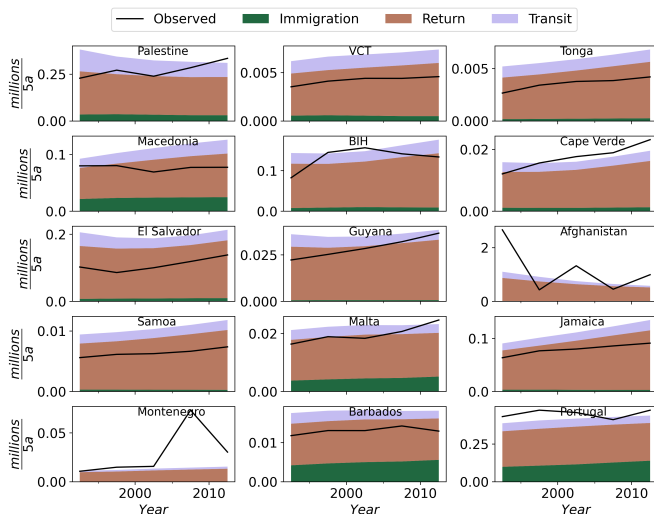
⁴In our modeling study, we focus on average incomes as represented by per-capita GDP, for simplicity and comparability with previous studies. It is worth noting, however, that there are other factors besides incomes that typically change as a country develops, and that also influence emigration trends; such as education or age structure (Skeldon 2008; Czaika and de Haas 2012). Replacing GDP by more comprehensive indicators of development, such as the Human Development Index (HDI), may be a promising direction for future work.

Figure 7: Bilateral migration flows



Note: Directed bilateral links are sorted in descending order by their average observed flow. Black lines show observed flows, and colors show simulations for the three different migration types (emigration from CoB, return, transit), stacked on top of each other such that the upper end of the colored area is the simulated total flow.

Figure 8: Countries and territories with large shares of simulated transit migration inflows



Note: "Immigration" denotes inflows of migrants from their respective countries of birth. Black lines show observed total inflows. VCT: Saint Vincent and the Grenadines; BIH: Bosnia and Herzegovina.

It must be noted again that we consider here the major underlying forces of migration in the long run. Shocks, such as through conflicts or disasters that cause refugee movements, or through political or economic crises or other events, are not accounted for in the migration model; nor are the effects of immigration policies and changes therein (Ortega and Peri 2013; Helbling and Leblang 2019). At the same time, the economic scenarios for the SSPs may be on the optimistic side, because the economic models, too, neglect the possibility of short-term shocks in the future that might affect growth rates (Dellink et al. 2017). Overall, this means that migration flows in our simulations may be at the lower end of what might actually materialize if the world followed any of the SSP storylines. The simulations should nonetheless be useful for evaluating the potential effects of different trajectories in long-run socioeconomic development.

Our study goes beyond previous projection exercises by explicitly accounting for the non-monotonic effects of origin-country incomes, diasporas, and the roles of transit and return migration; all of which have a substantial bearing either on the projected flows themselves or on the granularity at which they can be investigated. To achieve this, we have considered separate migration 'channels' for each place of birth. Nonetheless, we

neglect many other heterogeneities that influence migration. For instance, migration rates have been shown to differ by educational attainment, age, and sex, to varying degrees. Changes in educational attainment have been considered only implicitly here, through their effects on natural population growth and on economic development. Thus, a natural extension of our model would be to account separately for low-skilled and high-skilled migration (Docquier 2018).

Another possible extension is to account more explicitly for the rate of assimilation of migrants into the host society. We have assumed that first-generation immigrants are part of the diaspora for the rest of their lives, while their children are never part of the diaspora. In reality, first-generation immigrants may, over time, lose connections with their country of birth and thus, become irrelevant for the diaspora effect; and conversely, children of immigrants may have a substantial attachment to the birthplace of their parents, and may thus be able to assist or attract new immigrants. However, the rates of change in diaspora size that result from our simple assumptions are in a similar range as those resulting from empirical studies of migrant acculturation, which suggest that second-generation migrants have adapted, on average, about half way towards the host society's cultural values (Mesoudi 2018).

Finally, our assumptions about natural fertility and mortality rates of migrant populations are similarly simplistic, but our results are robust against alternative choices (Appendix, Figures A-5 and A-15). Overall, while uncertainties in both model parameters and assumptions are partly substantial, they still allow us to clearly separate the different SSPs and to draw robust conclusions about the qualitative behavior and approximate magnitude of future migration patterns, conditional on the GDP and natural population change scenarios and on our simplified global model design. Beyond the direct relevance of our projections for informing discussions of future global migration and its impacts, they lend themselves to inclusion into population models (KC and Lutz 2014; Jones and O'Neill 2016), where a more process-based treatment of migration could increase the realism of overall population projections.

5. Acknowledgments

The work was supported within the framework of the Leibniz Competition (K36/2017 IMPETUS) as well as the European Union Horizon 2020 programme (FUME) and the ERA4CS Joint Call on Researching and Advancing Climate Services (ISIpedia, BMBF grant number 01LS1711A). The authors thank Katja Frieler, Anders Levermann, and Bryan Jones for inspiring discussions.

References

- Abel, G.J. (2018). Non-zero trajectories for long-run net migration assumptions in global population projection models. *Demographic Research* 38(1): 1635–1662. doi:10.4054/DemRes.2018.38.54.
- Abel, G.J. and Cohen, J.E. (2019). Bilateral international migration flow estimates for 200 countries. *Scientific Data* 6(82): 1–13. doi:10.1038/s41597-019-0089-3.
- Azose, J.J. and Raftery, A.E. (2015). Bayesian probabilistic projection of international migration. *Demography* 52(5): 1627–1650. doi:10.1007/s13524-015-0415-0.
- Azose, J.J. and Raftery, A.E. (2019). Estimation of emigration, return migration, and transit migration between all pairs of countries. *Proceedings of the National Academy of Sciences* 116(1): 116–122. doi:10.1073/pnas.1722334116.
- Azose, J.J., Ševčíková, H., and Raftery, A.E. (2016). Probabilistic population projections with migration uncertainty. *Proceedings of the National Academy of Sciences* 113(23): 6460–6465. doi:10.1073/pnas.1606119113.
- Battistella, G. (2018). Return migration: A conceptual and policy framework. In: Appleby, J.K. and Kerwin, D. (eds.). *2018 International migration policy report: Perspectives on the content and implementation of the global compact for safe, orderly, and regular migration*. New York: Scalabrini Migration Study Centers. doi:10.14240/internationalmigrationrpt2018.
- Beine, M. (2016). The role of networks for migration flows: An update. *International Journal of Manpower* 37(7): 1154–1171. doi:10.1108/IJM-01-2016-0013.
- Beine, M., Docquier, F., and Özden, Ç. (2011). Diasporas. *Journal of Development Economics* 95(1): 30–41. doi:10.1016/j.jdeveco.2009.11.004.
- Beine, M., Docquier, F., and Schiff, M. (2013). International migration, transfer of norms and home country fertility. *Canadian Journal of Economics/Revue canadienne d'économie* 46(4): 1406–1430. doi:10.1111/caje.12062.
- Bijak, J., Disney, G., Findlay, A.M., Forster, J.J., Smith, P.W., and Wiśniowski, A. (2019). Assessing time series models for forecasting international migration: Lessons from the United Kingdom. *Journal of Forecasting* 470–487. doi:10.1002/for.2576.
- Bijak, J. and Wiśniowski, A. (2010). Bayesian forecasting of immigration to selected European countries by using expert knowledge. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 173(4): 775–796. doi:10.1111/j.1467-985X.2009.00635.x.
- Bove, V. and Elia, L. (2017). Migration, diversity, and economic growth. *World Devel-*

- opment 89: 227–239. doi:10.1016/j.worlddev.2016.08.012.
- Buettner, T. and Muenz, R. (2018). Modeling alternative projections of international migration. Washington: World Bank (KNOMAD working paper 3).
- Clemens, M.A. (2014). Does development reduce migration? In: Lucas, R.E. (ed.). *International handbook on migration and economic development*. Edward Elgar Publishing: 152–185. doi:10.4337/9781782548072.00010.
- Collier, P. (2013). *Exodus: How migration is changing our world*. Oxford: Oxford University Press.
- Constant, A.F. (2020). *Time-space dynamics of return and circular migration*. Cham: Springer International Publishing. doi:10.1007/978-3-319-57365-6_107 – 1.
- Czaika, M. and de Haas, H. (2012). The role of internal and international relative deprivation in global migration. *Oxford Development Studies* 40(4): 423–442. doi:10.1080/13600818.2012.728581.
- De Beer, J. (1993). Forecast intervals of net migration: The case of the Netherlands. *Journal of Forecasting* 12(7): 585–599. doi:10.1002/for.3980120705.
- de Haas, H. (2007). Turning the tide? Why development will not stop migration. *Development and Change* 38(5): 819–841. doi:10.1111/j.1467-7660.2007.00435.x.
- de Haas, H. (2010). Migration and development: A theoretical perspective. *International Migration Review* 44(1): 227–264. doi:10.1111/j.1747-7379.2009.00804.x.
- Dellink, R., Chateau, J., Lanzi, E., and Magné, B. (2017). Long-term economic growth projections in the shared socioeconomic pathways. *Global Environmental Change* 42: 200–214. doi:10.1016/j.gloenvcha.2015.06.004.
- Docquier, F. (2018). Long-term trends in international migration: Lessons from macroeconomic model. *Economics and Business Review* 4(1): 3–15. doi:10.18559/ebr.2018.1.1.
- Dubuc, S. (2012). Immigration to the UK from high-fertility countries: Intergenerational adaptation and fertility convergence. *Population and Development Review* 38(2): 353–368. doi:10.1111/j.1728-4457.2012.00496.x.
- European Commission Joint Research Centre and JRC (2018). *Many more to come? Migration from and within Africa*. Luxembourg: Publications Office of the European Union.
- Feenstra, R.C., Inklaar, R., and Timmer, M.P. (2015). The next generation of the Penn World Table. *American Economic Review* 105(10): 3150–3182. doi:10.1257/aer.20130954.

- Geiger, T. (2018). Continuous national gross domestic product (GDP) time series for 195 countries: Past observations (1850–2005) harmonized with future projections according to the shared socio-economic pathways (2006–2100). *Earth System Science Data* 10(2): 847–856. doi:10.5194/essd-10-847-2018.
- Gómez-Herrera, E. (2013). Comparing alternative methods to estimate gravity models of bilateral trade. *Empirical Economics* 44(3): 1087–1111. doi:10.1007/s00181-012-0576-2.
- Hatton, T.J. and Williamson, J.G. (2005). What fundamentals drive world migration? In: Borjas, G. and Crisp, J. (eds.). *Poverty, international migration and asylum*. London: Palgrave Macmillan: 15–38.
- Helbling, M. and Leblang, D. (2019). Controlling immigration? How regulations affect migration flows. *European Journal of Political Research* 58(1): 248–269. doi:10.1111/1475-6765.12279.
- Jones, B. and O’Neill, B.C. (2016). Spatially explicit global population scenarios consistent with the Shared Socioeconomic Pathways. *Environmental Research Letters* 11(8): 084003. doi:10.1088/1748-9326/11/8/084003.
- KC, S. and Lutz, W. (2014). The human core of the shared socioeconomic pathways: Population scenarios by age, sex and level of education for all countries to 2100. *Global Environmental Change* 181–192. doi:10.1016/j.gloenvcha.2014.06.004.
- Lutz, W., Sanderson, W.C., and Scherbov, S. (1998). Expert-based probabilistic population projections. *Population and Development Review* 24: 139–155. doi:10.2307/2808054.
- Massey, D.S., Arango, J., Hugo, G., Kouaouci, A., Pellegrino, A., and Taylor, J.E. (1993). Theories of international migration: A review and appraisal. *Population and Development Review* 19(3): 431–466. doi:10.2307/2938462.
- Mayer, T. and Zignago, S. (2011). Notes on CEPII’s distances measures: The GeoDist database. *SSRN Electronic Journal* 3–47. doi:10.2139/ssrn.1994531.
- Mesoudi, A. (2018). Migration, acculturation, and the maintenance of between-group cultural variation. *PLOS ONE* 13(10): e0205573. doi:10.1371/journal.pone.0205573.
- Migali, S., Natale, F., Tintori, G., Kalantaryan, S., Grubanov-Boskovic, S., Scipioni, M., Farinosi, F., Cattaneo, C., Benandi, B., Follador, M., Bidoglio, G., McMahon, S., and Barbas, T. (2018). International migration drivers: A quantitative assessment of the structural factors shaping migration. Luxembourg: Publications Office of the European Union.
- Milewski, N. (2010). Immigrant fertility in West Germany: Is there a socialization effect

- in transitions to second and third births? *European Journal of Population / Revue européenne de Démographie* 26(3): 297–323. doi:10.1007/s10680-010-9211-0.
- O'Neill, B.C., Krieglger, E., Ebi, K.L., Kemp-Benedict, E., Riahi, K., Rothman, D.S., van Ruijven, B.J., van Vuuren, D.P., Birkmann, J., Kok, K., Levy, M., and Solecki, W. (2017). The roads ahead: Narratives for shared socioeconomic pathways describing world futures in the 21st century. *Global Environmental Change* 42: 169–180. doi:10.1016/j.gloenvcha.2015.01.004.
- Ortega, F. and Peri, G. (2013). The effect of income and immigration policies on international migration. *Migration Studies* 1(1): 47–74. doi:10.1093/migration/mns004.
- Raymer, J., Wiśniowski, A., Forster, J.J., Smith, P.W.F., and Bijak, J. (2013). Integrated modeling of European migration. *Journal of the American Statistical Association* 108(503): 801–819. doi:10.1080/01621459.2013.789435.
- Riahi, K., van Vuuren, D.P., Krieglger, E., Edmonds, J., O'Neill, B.C., Fujimori, S., Bauer, N., Calvin, K., Dellink, R., Fricko, O., Lutz, W., Popp, A., Cuaresma, J.C., KC, S., Leimbach, M., Jiang, L., Kram, T., Rao, S., Emmerling, J., Ebi, K., Hasegawa, T., Havlik, P., Humpenöder, F., Da Silva, L.A., Smith, S., Stehfest, E., Bosetti, V., Eom, J., Gernaat, D., Masui, T., Rogelj, J., Strefler, J., Drouet, L., Krey, V., Luderer, G., Harmsen, M., Takahashi, K., Baumstark, L., Doelman, J.C., Kainuma, M., Klimont, Z., Marangoni, G., Lotze-Campen, H., Obersteiner, M., Tabeau, A., and Tavoni, M. (2017). The shared socioeconomic pathways and their energy, land use, and greenhouse gas emissions implications: An overview. *Global Environmental Change* 42: 153–168. doi:10.1016/j.gloenvcha.2016.05.009.
- Santos Silva, J.M.C. and Tenreyro, S. (2006). The log of gravity. *Review of Economics and Statistics* 88(4): 641–658. doi:10.1162/rest.88.4.641.
- Schmid, S. and Kohls, M. (2010). Reproductive behaviour of migrant women in Germany: Data, patterns and determinants. *Vienna Yearbook of Population Research* 2009: 39–61. doi:10.1553/populationyearbook2009s39.
- Skeldon, R. (2008). International migration as a tool in development policy: A passing phase? *Population and Development Review* 34(1): 1–18. doi:10.1111/j.1728-4457.2008.00203.x.
- UNHCR (2020). Refugee population statistics database [electronic resource]. <https://www.unhcr.org/refugee-statistics/download/>.
- United Nations (2015). International migration flows to and from selected countries: The 2015 revision [electronic resource]. New York: Department of Economic and Social Affairs, Population Division.

<https://www.un.org/en/development/desa/population/migration/data/empirical2/migrationflows.asp>

United Nations (2017). Trends in international migrant stock: The 2017 revision [electronic resource]. New York: Department of Economic and Social Affairs, Population Division. <https://www.un.org/en/development/desa/population/migration/data/estimates2/estimates17.asp>.

United Nations (2019a). International migrant stock 2019 [electronic resource]. New York: Department of Economic and Social Affairs, Population Division. <https://www.un.org/en/development/desa/population/migration/data/estimates2/estimates19.asp>.

United Nations (2019b). World population prospects 2019 [electronic resource]. New York: Department of Economic and Social Affairs, Population Division. <https://population.un.org/wpp/Download/Standard/Population/>.

Wesselbaum, D. and Aburn, A. (2019). Gone with the wind: International migration. *Global and Planetary Change* 178: 96–109. doi:10.1016/j.gloplacha.2019.04.008.

Willekens, F., Massey, D., Raymer, J., and Beauchemin, C. (2016). International migration under the microscope. *Science* 352(6288): 897–899. doi:10.1126/science.aaf6545.

Williamson, J.G. (2015). World migration in historical perspective. *Handbook of the Economics of International Migration* 1: 89–101. doi:10.1016/B978-0-444-53764-5.09983-7.

Wolf, K. and Mulder, C.H. (2019). Comparing the fertility of Ghanaian migrants in Europe with nonmigrants in Ghana. *Population, Space and Place* 25(2): e2171. doi:10.1002/psp.2171.

Appendix: Global bilateral migration projections accounting for diasporas, transit and return flows, and poverty constraints

A. Model coverage

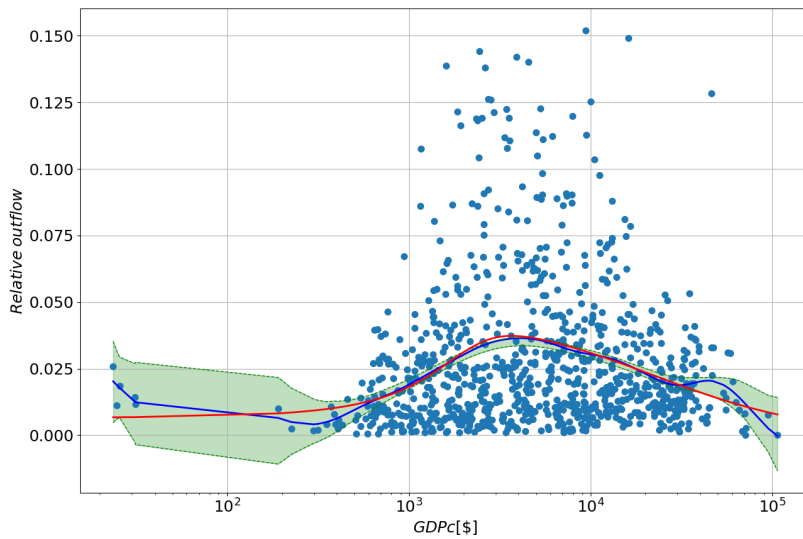
The following countries and territories (in descending order by population in 2015) are not included in the dynamic model, due to missing data:

- Sudan, Dem. People's Republic of Korea, South Sudan, Timor-Leste, Réunion, Micronesia, Guadeloupe, Martinique, China Macao SAR, Western Sahara, French Polynesia, New Caledonia, Curaçao, Channel Islands, Guam, French Guiana, United States Virgin Islands, Grenada, Mayotte, Kiribati, Dominica, Seychelles, Isle of Man, Antigua and Barbuda, Aruba, Bermuda, Greenland, Andorra, Faeroe Islands, Marshall Islands, American Samoa, Northern Mariana Islands, Saint Kitts and Nevis, Monaco, Gibraltar, Liechtenstein, Sint Maarten (Dutch part), Cayman Islands, San Marino, Cook Islands, British Virgin Islands, Palau, Wallis and Futuna Islands, Turks and Caicos Islands, Montserrat, Nauru, Tuvalu, Anguilla, Saint Pierre and Miquelon, Saint Helena, Niue, Falkland Islands (Malvinas), Tokelau, Holy See.

Table A-1: The 30 largest bilateral migration flows in the A19 dataset for which origin country GDP per capita is more than twice as high as destination GDP per capita. Flows are cumulated over 1990–2015, while the GDP ratio is the average over the same period

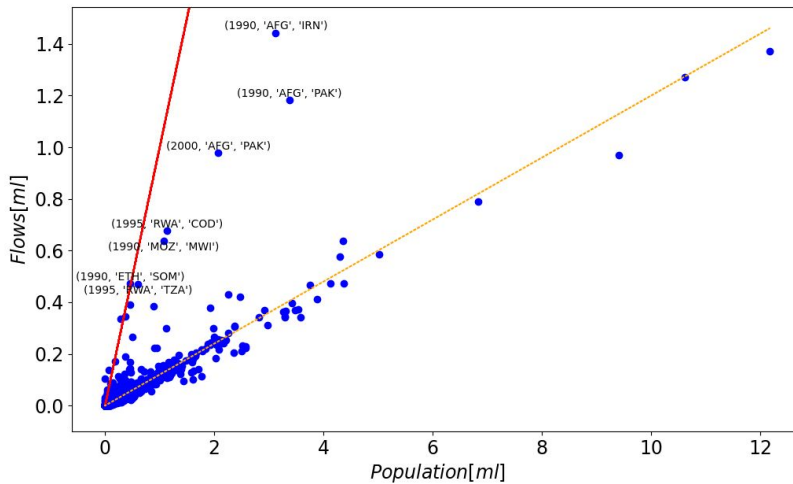
Origin	Destination	Flow (millions)	GDP ratio (original destinations)
USA	MEX	5.97	3.6
RUS	UKR	3.20	2.3
IRN	AFG	1.99	9.2
PAK	AFG	1.92	2.7
HKG	CHN	1.50	7.6
RUS	UZB	1.17	3.2
DEU	TUR	1.11	2.7
DEU	RUS	1.04	2.6
ARE	IND	0.83	33.2
USA	PRI	0.83	2.7
USA	CHN	0.83	9.3
USA	PHL	0.81	10.2
SAU	IND	0.75	9.8
CIV	BFA	0.68	2.5
FRA	DZA	0.62	4.1
USA	IND	0.60	20.0
USA	RUS	0.55	3.6
USA	VNM	0.55	20.3
SAU	YEM	0.53	16.9
USA	CUB	0.51	9.3
SAU	IDN	0.51	5.5
DEU	KAZ	0.50	3.7
SGP	MYS	0.47	3.1
THA	KHM	0.46	5.4
THA	MMR	0.46	8.1
FRA	MAR	0.46	7.0
SAU	PAK	0.46	8.5
SDN	ETH	0.45	3.3
RUS	KAZ	0.45	2.0
GIN	LBR	0.44	4.9

Figure A-1: Gross emigration rates versus origin GDP per capita, as shown in Figure 1 in the main paper, but including a non-parametric, local-linear regression with a Gaussian kernel (blue line)



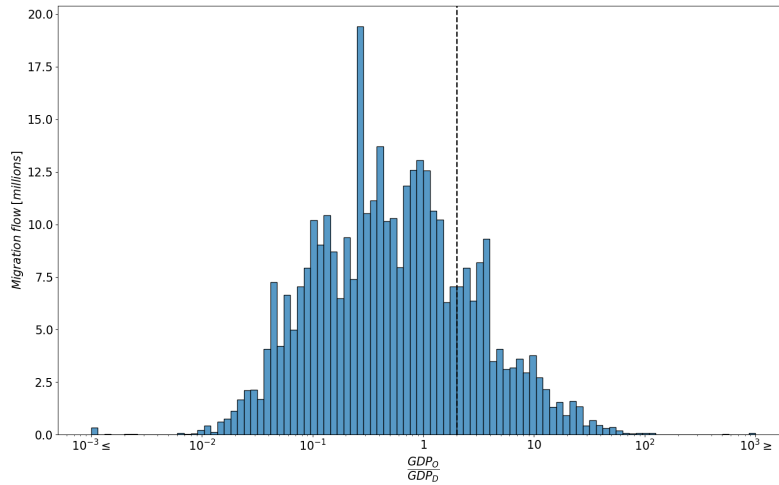
Note: A confidence interval of 66%, obtained through bootstrapping, is shown by the green colored area. The red line shows a fit of the function $F(G_i)$ Equation 2 in the main paper), and is identical to the green line in Figure 1 in the main paper.

Figure A-2: Return migration flows by Azose and Raftery (2019) and Abel and Cohen (2019) versus corresponding bilateral migrant stocks. Red line has slope 1; dashed yellow line has slope 0.12, i.e., the value of the proportionality factor b in our return migration model, Equation 3 (main paper)



Note: The most apparent outliers are labeled with the first year of the period, the origin country, and the destination country (i.e. country of birth). All of them are flows from countries with ongoing civil wars, and/or regime changes, at the time into neighboring countries, thus plausibly exceeding “normal” return flow rates. Note that these CoB-specific flow estimates are used here only for illustrative purposes, and not used elsewhere in our study.

Figure A-3: Size of bilateral flows from the A19 dataset versus GDP per capita ratio between origin and destination country



Note: Dashed vertical line indicates a ratio of 2; flows to the right of this line are assumed to be predominantly return flows, and are removed from the dataset for the estimation of Equation 2 in the main paper.

Figure A-5: Regional historical population growth, as observed (black) and as simulated assuming diasporas growing either at the same rate as their country of birth (green) or at the same rate as their residence country (red)

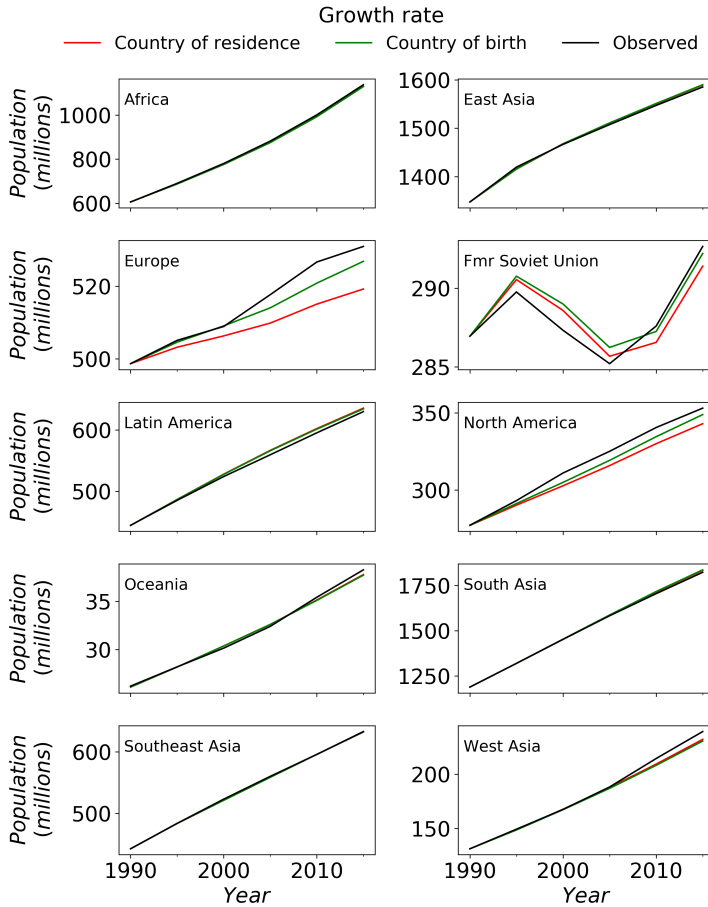
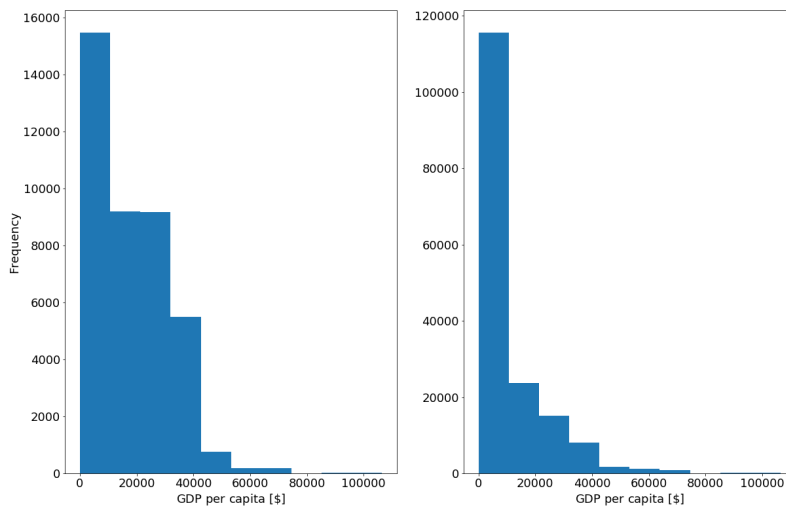


Figure A-6: Distribution of destination country GDP per capita in the bilateral links included in the alternative flow dataset derived through stock differencing (left) and in the main A19 flow dataset (right)



Note: Note different y-axis scales.

Figure A-7: Effect of parameter uncertainty on model simulations of net migration in an exemplary region (Africa, left) and country (Germany, right)

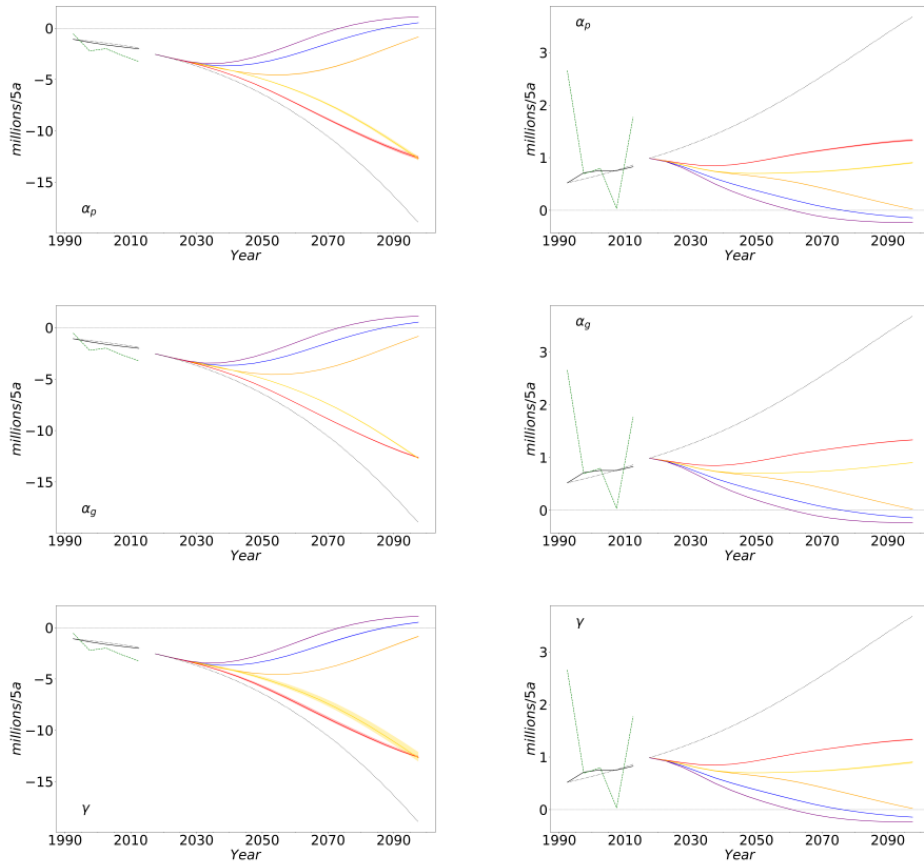
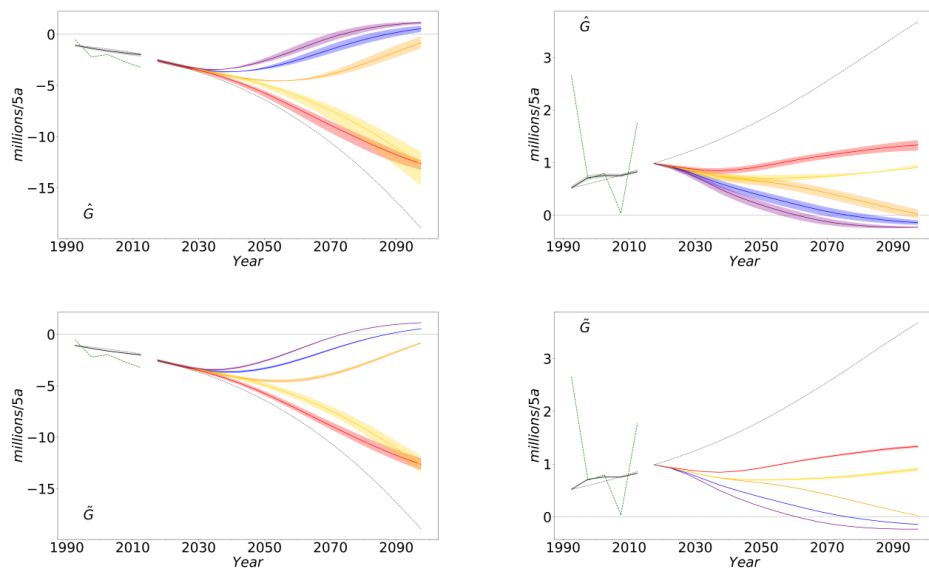
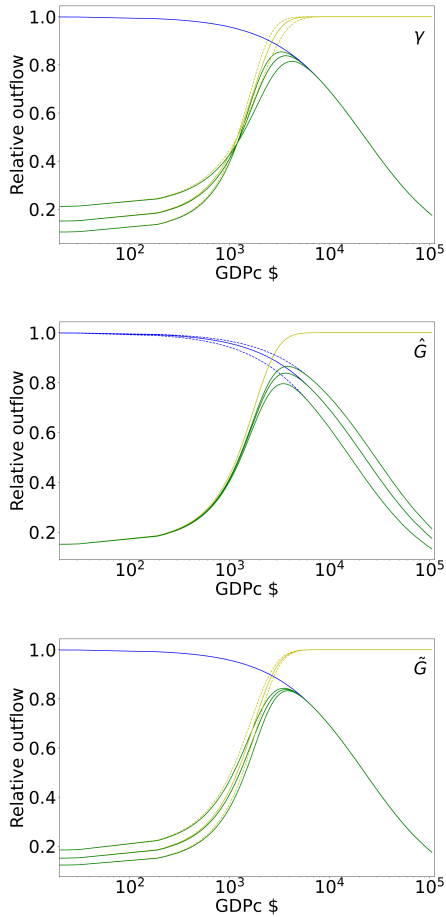


Figure A-7: Continued



Note: Solid lines are identical to those in the corresponding panels of Figure 4 and 5 in the main paper. Each shaded band represents a pair of additional simulations where one parameter at a time is set to the upper and lower bound, respectively, of its 66% confidence interval; while all other parameters are held at their central values, as reported in Table 1 in the main paper. Parameters under variation are, from top to bottom, α_p , α_g , γ , \bar{G} , \vec{G} .

Figure A-8: Illustration of the effect of parameter uncertainty on the shape of $F(G_i)$ (green) and its two components, $F_{intent}(G_i)$ (blue) and $F_{resource}(G_i)$ (yellow; cf. Equation 2 in the main paper)



Note: In each panel, one of the three parameters of equation 2 is varied between its central estimate (solid line) and the upper and lower bound of its 66% confidence interval (dotted lines); while the other two parameters are held at their central values, as reported in Table 1 in the main paper. Parameters under variation are, from top to bottom, γ , \hat{G} , \tilde{G} .

Figure A-9: As Figure 4 in the main paper, but expressed as percentage of the total population in each region and timestep; and omitting the shading

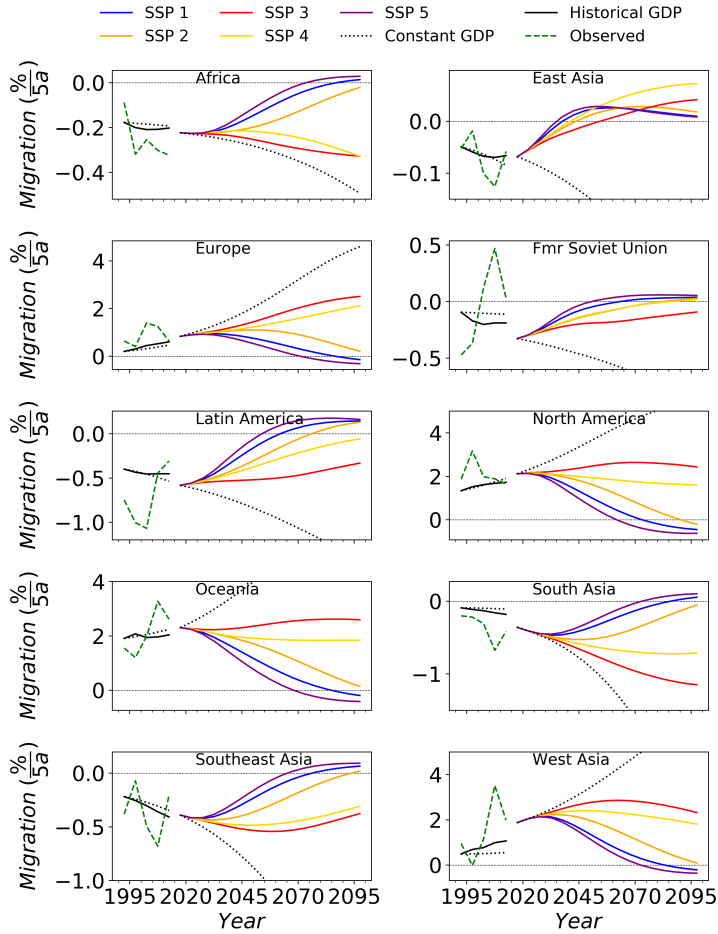


Figure A-10: As Figure 5 in the main paper, but expressed as percentage of the total population in each country and time step; and omitting the shading

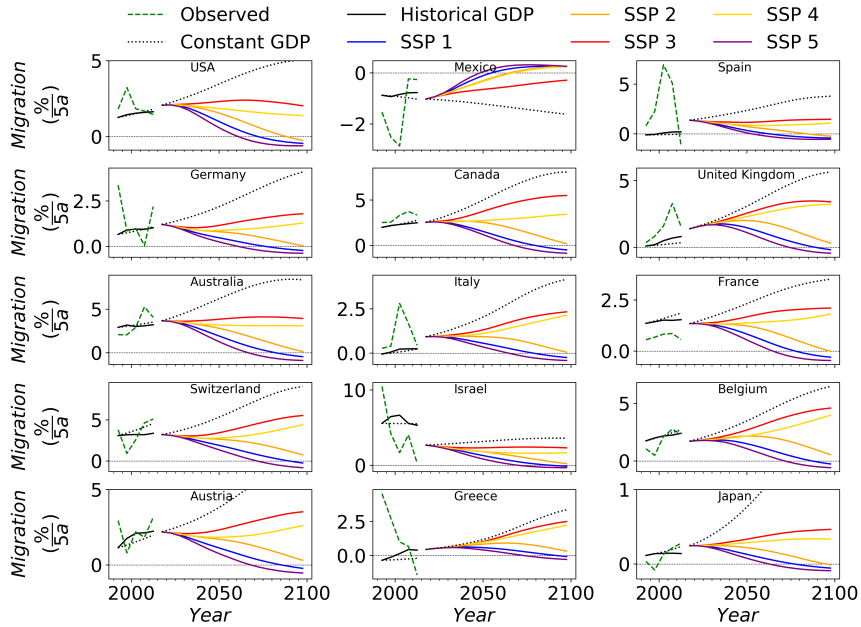


Figure A-11: As Figure 6 in the main paper, but expressed as percentage of the total population in each country and time step; and omitting the shading

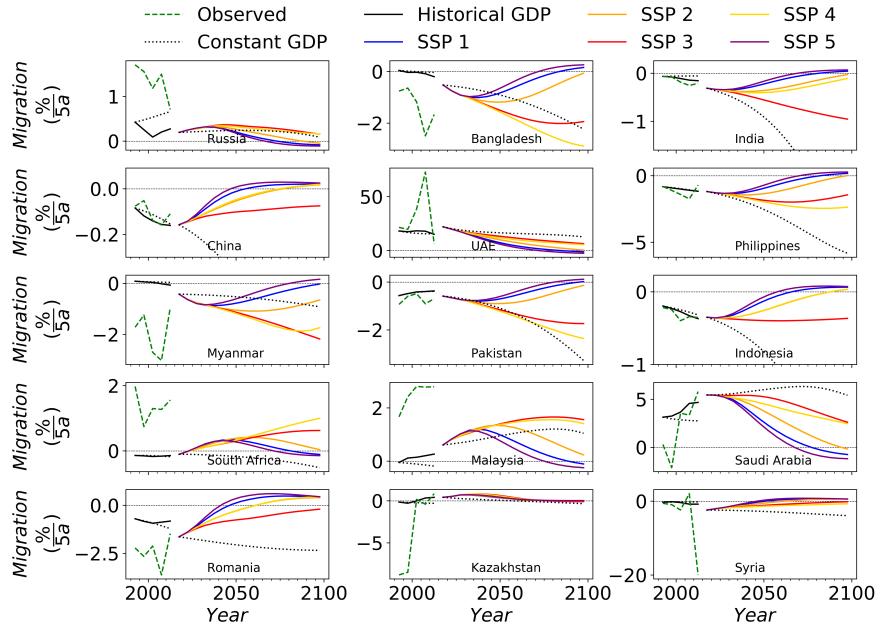


Figure A-12: As Figure 8 in the main paper but for gross outflows

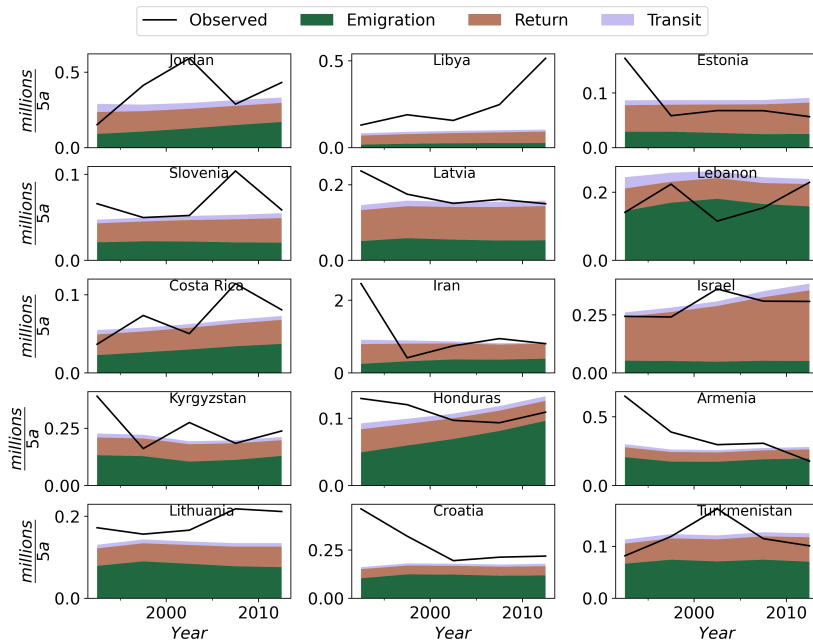
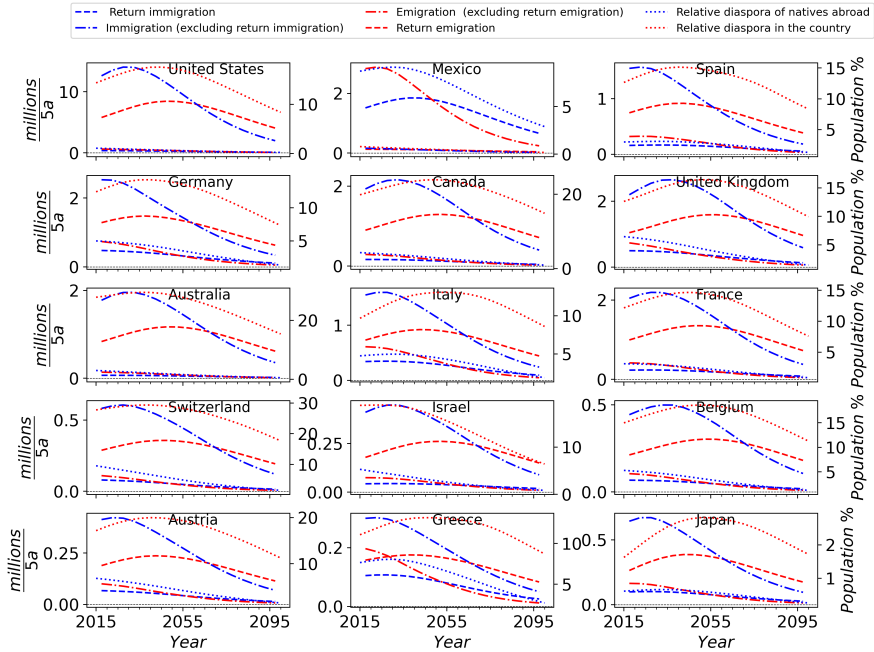


Figure A-13: Role of return migration flows in country-level net migration, for SSP 5 and the countries shown in Figure 5 of the main paper



Note: Dashed and dash-dotted lines (left vertical axis) show the return flow and residual components, respectively, of immigration into (blue) and emigration out of (red) each country. Dotted lines (right vertical axis) show the share of nationals living abroad (blue) and foreigners within the country (red).

Figure A-14: As Figure A-13 but for the countries shown in Figure 6 of the main paper.

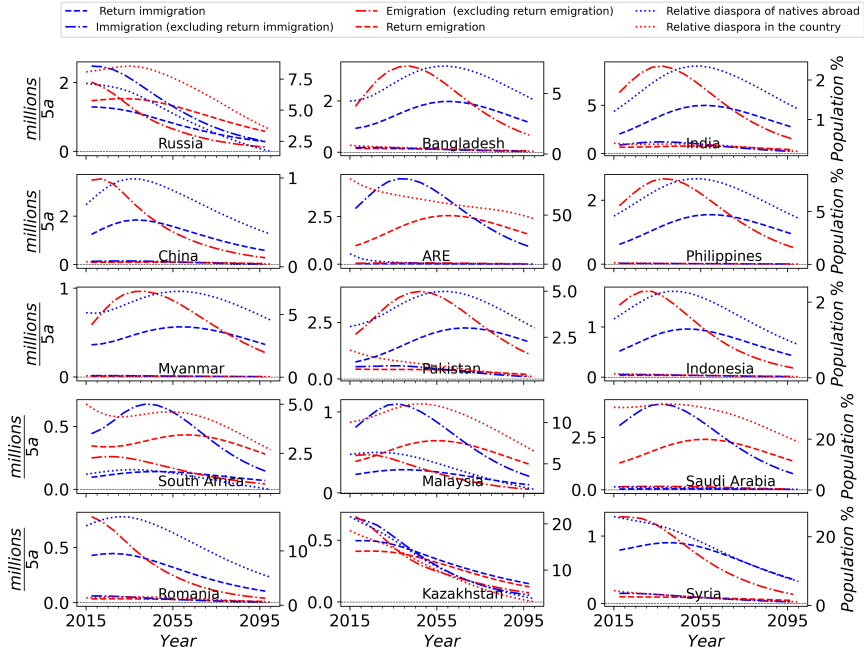


Figure A-15: Simulated regional net migration assuming diasporas growing either at the same rate as their origin country (dashed lines), or at the same rate as their residence country, as in the main paper (dotted lines)

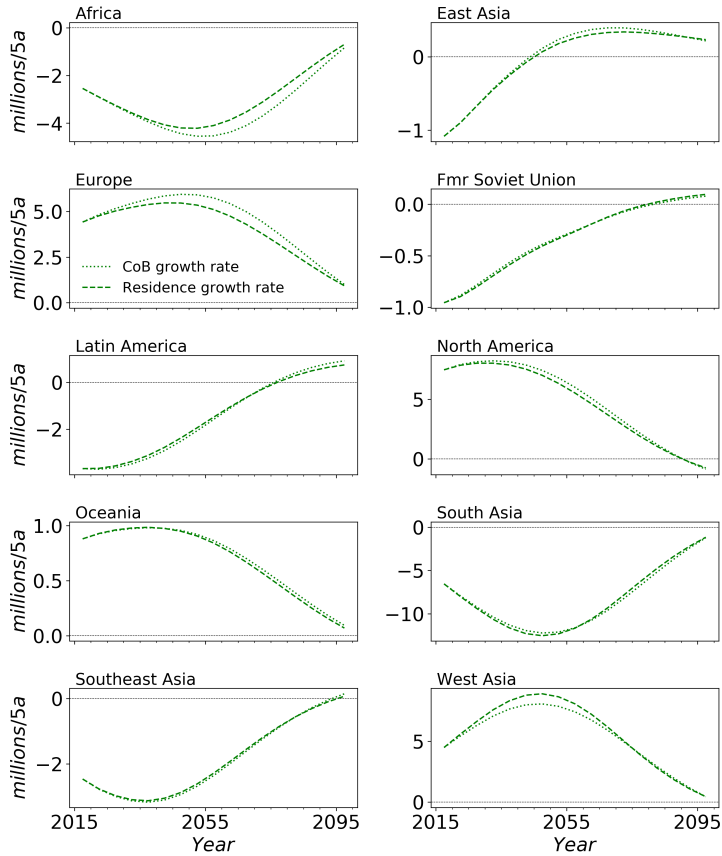


Figure A-16: As Figures A-10 and A-11 but for further countries in Africa

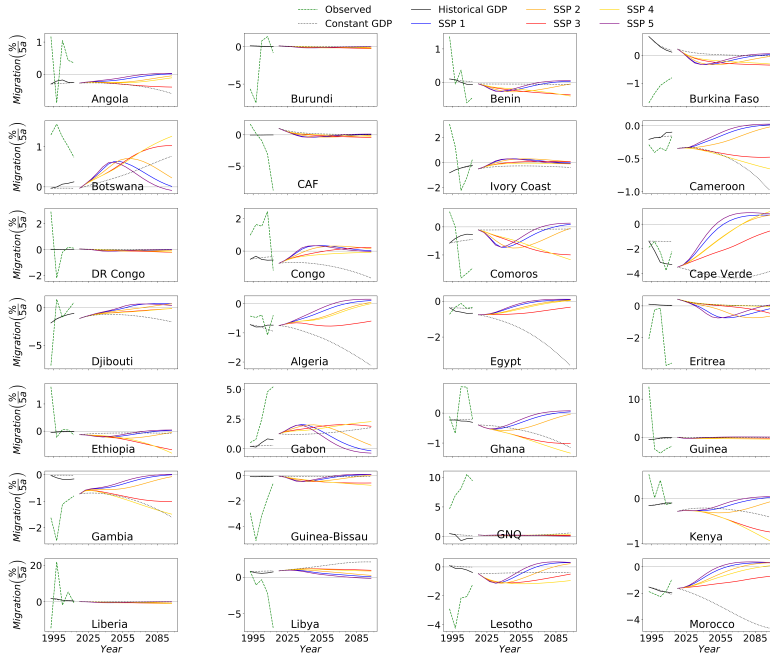


Figure A-16: Continued

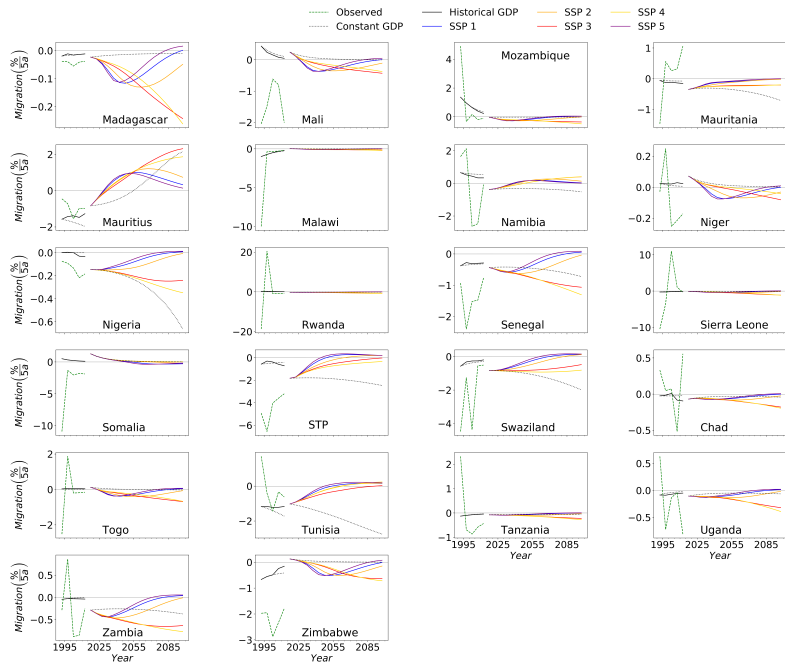


Figure A-17: As Figures A-10 and A-11 but for further countries in Asia

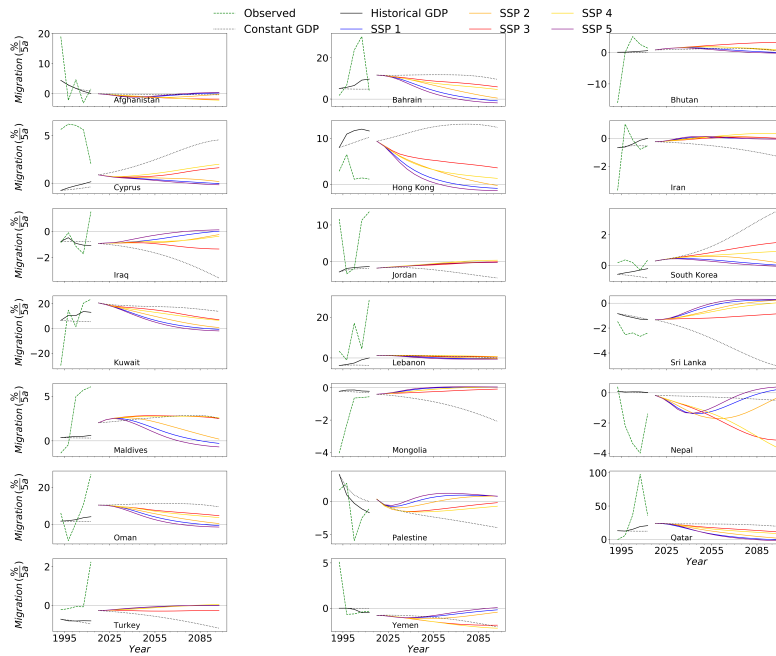


Figure A-18: As Figures A-10 and A-11 but for further countries in Europe

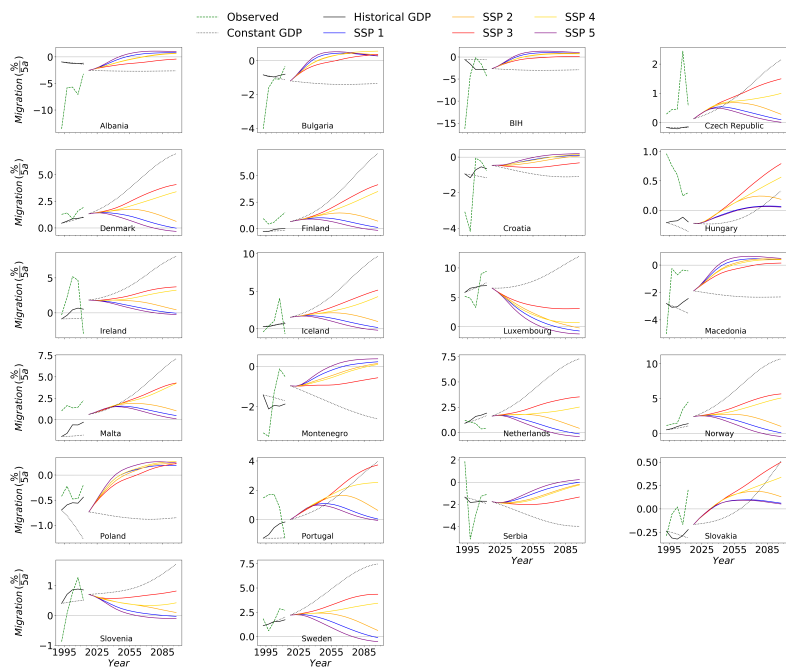


Figure A-19: As Figures A-10 and A-11 but for further countries of the Former Soviet Union

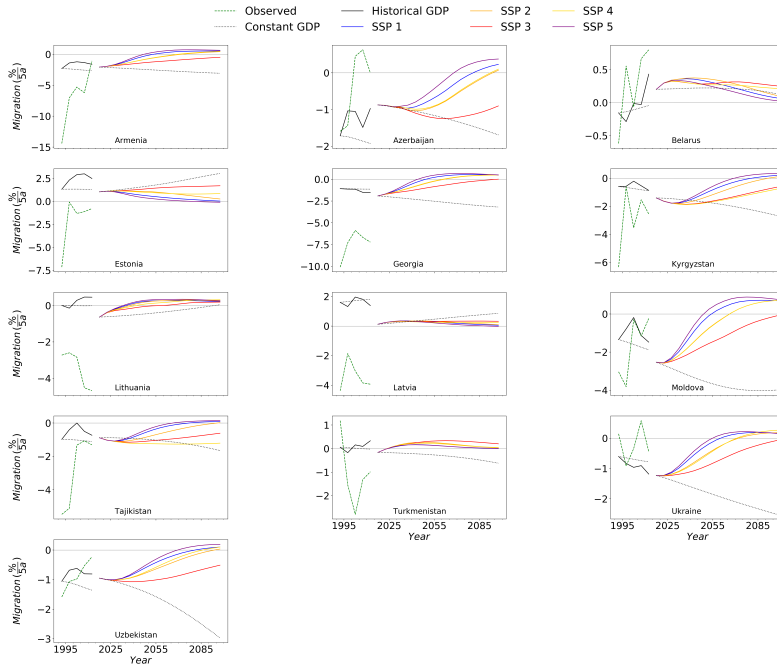


Figure A-20: As Figures A-10 and A-11 but for further countries in Latin America

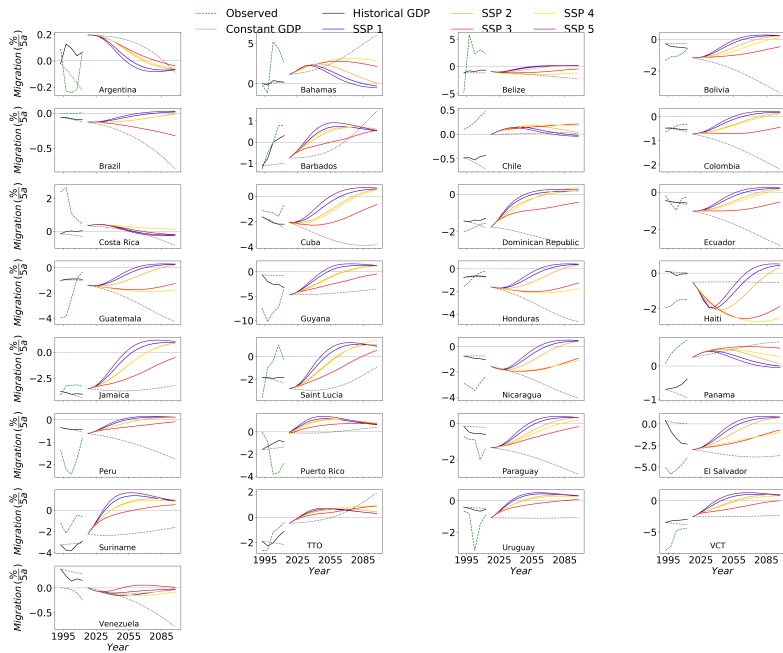


Figure A-21: As Figures A-10 and A-11 but for further countries in Oceania

