PNAS
www.pnas.org

Supplementary Information for

Population aging, migration, and productivity in Europe

Authors: Guillaume Marois[1,2]*, Alain Bélanger[2,3], Wolfgang Lutz[2]

[1] Asian Demographic Research Institute, Shanghai University, 99 Shangda Road, Baoshan, Shanghai, 200444, China

[2] Wittgenstein Centre for Demography and Global Human Capital (Univ. Vienna, IIASA, VID/ÖAW), International Institute for Applied Systems Analysis, 2361 Laxenburg, Austria.

[3]Centre Urbanisation Culture Société, Institut national de la recherche scientifique, 385 rue Sherbrooke Est, Montréal, Canada H2X 1E3.

***Correspondence to**: marois@iiasa.ac.at
ORCID: https://orcid.org/0000-0002-2701-628

This PDF file includes:

SI Appendix 1 – Data, methods, assumptions, and scenarios
SI Appendix 2 – Detailed results by scenarios and country, 2015, 2030 and 2060
SI References

www.pnas.org/cgi/doi/10.1073/pnas.1918988117

Supplementary Information Text

Appendix 1 – Data, methods, assumptions, and scenarios

This appendix presents in more detail the microsimulation model developed to realize population projections. It also presents the scenarios and their assumptions and provides a definition of the indicators used in the paper, in particular the productivity-weighted labor force dependency ratio.

The projection model: CEPAM-Mic

The population projections are realized using a microsimulation model called CEPAM-Mic that allows the study of alternative scenarios and their consequences for future population trends in the European Union. The model was built as part of the Centre of Expertise on Population and Migration (CEPAM) project, a partnership between the Joint Research Centre (JRC) of the European Commission and the International Institute for Applied Systems Analysis (IIASA). This collaborative effort has been designed to assess the longer-term implications of migration- and population-related challenges.

Microsimulation is an alternative approach to the deterministic macro-level population projection models that use aggregate level data to project future population dynamics. In microsimulation the modelling is based on individual level data. The baseline population consists of individual actors whose individual characteristics represent the composition of a given population across chosen dimensions. These individual actors are exposed to the risk of a set of events relevant to their state and specific to their own characteristics – death, births of a child (which generates a new actor inside the model), moving to a different EU country, leaving to a non-EU country, achieving next level of education, entering or exiting the labor market and so on. Immigrants from non-EU countries enter the model with a predetermined set of individual characteristics and are subjected to risk of the events mentioned above. The population is simulated in continuous time and the transitions between the states are determined stochastically. When a transition occurs, the characteristics of the individual are changed and the risks are recalculated. Dynamic microsimulation in continuous time thus allows for not only including a larger set of dimensions than the standard multistate population projection models in which handling more than three or four dimensions becomes challenging, but also for handle competing risks easily.

CEPAM-Mic allows for the study of alternative scenarios of migration and their consequences on future population and labor supply trends in the European Union. It is developed using the Modgen language, which is a microsimulation programming language developed by Statistics Canada, integrated into the Microsoft Visual Studio C++ environment (1). The model is built following the framework proposed by Belanger et al. (2) to study population changes in a context of relatively high immigration and low fertility. CEPAM-Mic can thus dynamically project the population for EU28 member states under several socioeconomic and ethnocultural dimensions. Its base population counts 13 variables:

- age,
- sex,
- country of residence,
- student status,
- labor force participation,
- employment

- age at immigration,
- region of birth (11 clusters of world countries),
- duration of residence (4 categories),
- education level (3 categories),
- education of the mother (3 categories),
- religion (4 categories),
- language (3 categories).

The use of microsimulation becomes necessary given the large number of categories in each of these dimensions or variables. Indeed, microsimulation is a powerful tool that can replace traditional multistate projections when the number of dimensions becomes large (3) and also allows for the use of complex statistical models to project life-course transitions and events. This microsimulation model is characterized by the stochastic simulation of individual life courses. Simultaneous simulation of individual life courses allows the model to dynamically update the risks of various events based on an individual's state values, and further allows interactions between actors. CEPAM-mic model utilizes the interaction to model intergenerational transmission of education by linking child's characteristics to those of the mother.

Microsimulation methods are also very flexible in the sense that they allow for the creation of scenarios[1] combining different hypotheses concerning the future evolution of the stochastic parameters that drive the component of population changes (3, 4).

A detailed description of the different modules that make up this program is available elsewhere (2, 5–7). However, a short description of the different components of the model is useful, starting with the base population.

<u>The base population</u>

As mentioned previously, the baseline population is structured by 13 variables. However, there is no single data source (a survey or a census microdata files) that would contain all these variables and for all 28 EU member state. Therefore, we could not create the base population stock from a single data source, but we had to triangulate several surveys and integrate the information from them.

The main data source is European Labour Force Survey (LFS) as it includes all projected variables, except for religion and language spoken at home. For this reason and also the fact that its size is relatively large compared to other surveys, LFS is used as the main data source to build the base population and also to estimate several parameters of the microsimulation model. Although relatively large, the sample sizes of the LFS vary from country to country. Given the importance of analyzing the behavior of population groups which can sometimes be relatively small, it was useful to pool two years of the survey (2014 and 2015) to increase accuracy and reduce the Monte Carlo error[2]. Each record from this base population is an individual (an actor) in the microsimulation model (n=8,148,874). In addition, microdata from waves 1 to 7 of the European Social Survey (ESS) were used to impute the missing variables, religion and language spoken at

[1] Scenario is, in this case, defined as a combination of assumptions across the projection components (variables). Assumption specifies future value of a component or a parameter.

[2] The Monte Carlo error is the uncertainty resulting from the stochastic process of the microsimulation. The smaller is the N, the higher is the error.

home, in the base population. Imputation was done using polytomous logistic regressions in the mice package in R (8).

Compared to census data, LFS seems to measure some variables less well. In particular, the immigrant population of some countries seems to be significantly underestimated, which led us to calibrate the LFS data using European census data and other sources. The microdata set is thus calibrated in three steps: in the first step, the base population is reweighted to match the 2011 European census by country, age, sex, educational attainment, and place of birth (where available); in the second step, the resulting population from step 1 is adjusted to match the religion distribution by country and sex (9); the third and final step occurs during the projection where the population is calibrated on the 2015 Wittgenstein center's estimates by age, sex, education, and country (10).

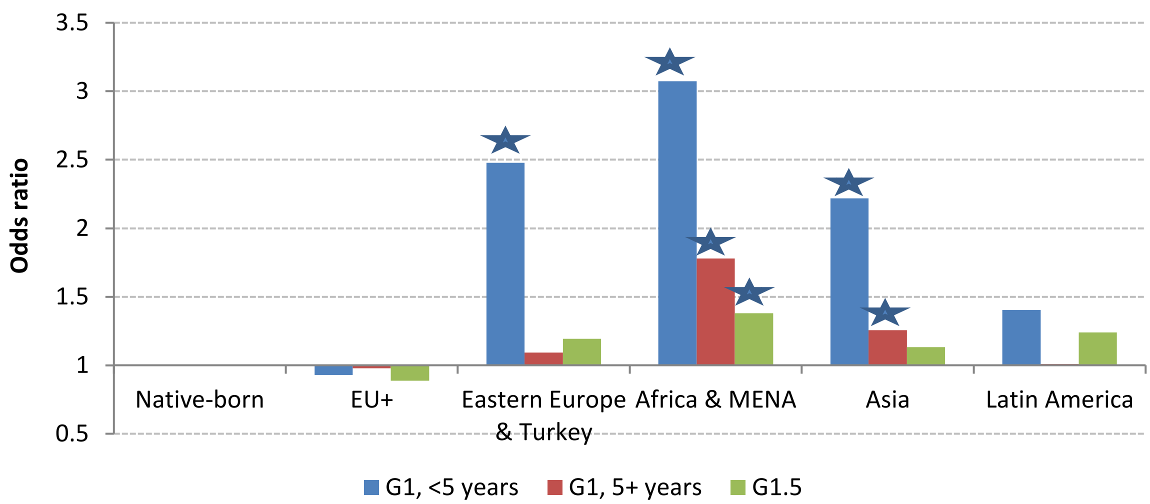
Mortality module and assumptions

Mortality rates by age, sex and educational attainment are taken from Lutz et al. (11). Future trends for these rates were set combining statistical models with expert judgment (12). Continuous improvement in life expectancy is assumed with long-term regional convergence, exceeding 90 years in most European countries by 2060. Differentials in life expectancy (at the age of 15) between the low and the high educated is, however, kept at about 4 years for females and 6 years for males.

Fertility module and assumptions

Using the own-children method with LFS data, differentials by age, country, education, student status (with an interaction with age), region of birth, age at immigration, and duration of stay were estimated using logit regression models (7). Outcomes (see figure S1) show higher fertility for immigrants (generation 1) from some regions such as Sub-Saharan countries, Middle East, and North Africa (MENA). The fertility is also higher for recent immigrants, but tends to converge with natives with duration of stay. For immigrants arrived during childhood (generation 1.5), fertility levels fall between that of their parents and of the natives.

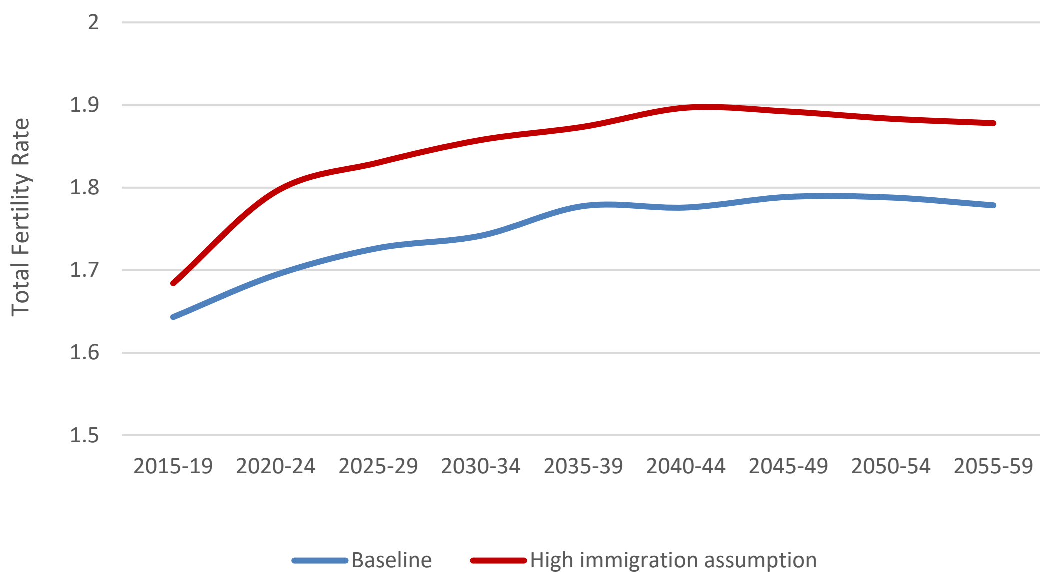
Fig. S1: Odds ratios of giving birth by region of origin, generational status and duration of stay, native-born = ref



Note: Statistically significant results marked by asterisk
Source: Potancoková and Marois (7)

An additional parameter for the student status reduces fertility of women who are still in school (OR=0.123). Parameters for the student status and the immigration variable are then contrasted to the weighted population average. These adjusted parameters are then added to the base age-, education-, and country-specific fertility rates from Lutz et al. (11), which values and trends were determined after a large expert survey in the field of fertility studies (13). Experts assumed slightly increasing fertility for a majority of EU countries due to the ongoing process of fertility postponement. These assumptions on overall fertility level in the EU countries in hand with the changing population composition by educational attainment (increasing due to the assumed education expansion) and nativity status (see migration assumptions below) and at current differentials result in an increased total fertility rate from 1.6 in 2015-19 to 1.8 in 2055-59 (see figure S2) for the whole EU28 for the baseline scenario and 1.9 for scenarios with high immigration.

Fig. S2: projected total fertility rates (TFR) for EU28



Source: Potancoková and Marois (7)

Education

The education module of CEPAM-Mic is exhaustively described in (6). In summary, CEPAM-Mic includes three levels of education, either:

(1) Low: Lower secondary or less (ISCED 1 and 2);
(2) Medium: Upper secondary completed (ISCED 3);
(3) High: Postsecondary (ISCED 4+).

The highest level of education that an individual will reach over the life course is set probabilistically at birth (or at arrival for immigrants who arrived during childhood) using parameters estimated from an ordered logit regression. The model explicitly considers the influence of personal characteristics and the education of the mother. Sex- and country-specific cohort parameters are also included and extrapolated to establish assumptions for future cohorts. The model equation is thus formulated as follows:

Eq.S5 $$\ln\left(\frac{E_{ij}}{1-E_{ij}}\right) = \beta_{0j} + \beta_{1j}Ct_i + \beta_{2j}Cr_i + \beta_{3j}(Ct_i * Cr_i) + \beta_{4j}X_i + \beta_{5j}Z_i$$

Where

- E_{ij} is the probability that an individual i reaches level of education j, where j equals High or Medium;
- Ct is the country;
- Cr is a discrete variable for cohorts (1940-44=1; 1945-49=2, …, 1975-1979=8);
- X is a set of sociocultural variables;
- Z is the education of the mother.

Sociocultural variables include language, religion, and place of birth. As such, the education module implements differentials in the educational pathways for children with an immigrant background as well as for different social classes as reflected by the education of the mother. Parameters show that the mother's education is a strong predictor of children's future educational attainment, while sociocultural variables, such as being Muslim (especially for women) or speaking a non-European language at home decrease the odds of getting postsecondary education. These parameters are consistent with the existing literature trajectory (14–17). In all scenarios of this paper, parameters β_4 and β_5 are kept constant throughout the projection, while country-specific cohort trends (β_0 to β_3) are extrapolated over the time span of the projection (postsecondary is capped at 90%, as what has been done in other international projection of educational attainment (11, 18)).

Individuals are set as student starting from age 5 until the age of graduation from the highest completed level. The age at graduation is determined for all degrees using Eurostat distributions by ISCED levels for the latest graduated cohorts (2013-2014). The modeling of education thus allows distinguishing, for each individual at each projection step, the highest level of education that will be reached during the life course, the current level of education, and whether or not the individual is still in schooling.

Migration

To allow greater flexibility in the development of migration assumptions, international immigration, which in the context of this research corresponds to that from non-European countries (i.e. the third countries), is treated separately from other migration movements, either international emigration to outside the EU or migration between EU countries (termed intra-EU mobility in this report[3]).

Out-migration parameters were estimated in three steps. First, the intensity of out-migration is determined by computing country-level out-migration rates by sex and country of residence from the average number of out-migrants between 2013 and 2016 (Eurostat table: migr_emi2) divided by the average population aged 20-34 during the same period. The age group 20-34 was chosen as the exposure for the rate because majority of outmigrants are in that age at migration. Using whole population as an exposure yields distorted rate(19). In a second step, age-specific out-migration rates are derived within the microsimulation model as follows: first, country-level out-migration rates are applied to the projected population aged 20-34 to get the expected number of out-migrants in a given period. The number of out-migrants is then distributed according to age using a Rogers-Castro age schedule (20). Finally, the age-specific out-migration rates are obtained by taking the ratio of out-migrants to the population, by age, sex and country of residence. Out-migration rates in the simulation are recalculated every five years.

During the simulation, out-migrants may either move within the EU, and are assigned a new country of residence, or they can leave the EU, in which case their simulation is terminated. The proportion of out-migrants leaving the EU is derived from Eurostat tables on emigration according to the region of destination (table: migr_emi3nxt). Origin-destination matrices for intra-European mobility were derived using updated estimates for the period 2009-2016 of Raymer et al.'s (21) Bayesian estimates of intra-EU mobility. Country-specific calibration factors are then calculated from a preliminary simulation for the period 2013-2016 in order to get the same number of

[3] The initial estimates pertain to moves of EU residents between the EU Member States.

entrances by country than what was estimated by Eurostat for the same period. These calibration factors are kept constant for the rest of the projection.

International immigration is treated as a separate component and the model allows great flexibility on assumptions about the future number of immigrants and their characteristics. Since the model does not project the rest of the world population, the size of future immigrant cohorts is an exogenous component and the annual number of immigrants by country is directly entered as a parameter[4]. In this paper, there are two assumptions concerning the number of immigrants. In all scenarios, destination countries in the EU are kept proportionally constant.

1. For the **baseline volume** assumption of CEPAM-Mic, the number of international immigrants is assumed to remain constant to average observed during the period 2013-2016. In order to correct the exceptionally high immigration inflows during the peak of the so-called refugee crisis in 2015-2016, we have adjusted the extreme values for Austria and Germany, and the flow of 2016 for Greece. The number of international immigrants settling in Europe in a 5-year period is thus assumed to be about 10M. This number of international immigrants is assumed to remain constant for all periods until the projection horizon in 2060. The immigrants are distributed into the EU Member states according to the respective shares in the 2013-2016 Eurostat data and the resulting overall flows are presented in Table S1.
2. The **high volume** assumption doubles the number of immigrants assumed in the baseline assumption. Under this variant, approximately 20 million international immigrants land in the EU every 5 years. Although much higher than what observed in Europe in the past, this assumption is still realistic, since proportionally to its population, it corresponds to the immigration rate observed in high immigration countries such as Canada in last decades (about 0.75%/year).
3. For the **low volume** assumption, the number of immigrants settling in the EU amounts to approximately 1.2 million every 5 years (corresponding to the recent immigration rate in Japan).

[4] On the long run, using fixed inflows for immigration and rates for emigration could impact net migration trends, as emigration flows may increase or decrease depending on the population size and structure.

Table S1: Assumptions on the number of international immigrants (born outside EU28) (5 years inflow)			
Host country	**Baseline volume**	**High volume**	**Low volume**
AT	223,597	447,193	26,850
BE	263,245	526,490	31,610
BG	84,909	169,818	10,196
CY	31,609	63,218	3,796
CZ	96,489	192,978	11,586
DE	1,991,155	3,982,310	239,094
DK	188,434	376,868	22,626
EE	16,409	32,818	1,970
ES	1,100,676	2,201,353	132,168
FI	85,108	170,215	10,220
FR	1,101,813	2,203,625	132,304
GR	99,462	198,923	11,944
HR	37,968	75,935	4,560
HU	114,329	228,658	13,728
IE	174,873	349,745	20,998
IT	1,057,411	2,114,823	126,972
LT	37,193	74,385	4,466
LU	9,348	18,695	1,122
LV	20,310	40,620	2,438
MT	34,791	69,583	4,178
NL	397,853	795,705	47,774
PL	466,226	932,453	55,984
PT	55,871	111,743	6,708
RO	147,265	294,530	17,684
SE	469,441	938,883	56,370
SI	51,148	102,295	6,142
SK	6,348	12,695	762
UK	1,630,214	3,260,428	195,754
EU28	9,993,490	19,986,979	1,200,004

Characteristics of recent immigrants in the base population are used as a basis to determine the characteristics of future immigrants in the simulation. Through reweighting, it is possible to change the immigrant distribution according to age, sex, education, place of birth, religion, and language in each country. This paper presents a set of three assumptions concerning the educational attainment of immigrants, summarized in table 2.

- The **baseline education assumption** supposes that future immigrants will have the same educational attainment than immigrants who arrived into the EU between 2001-2011 (estimated from the immigrants in the base population).
- The **high education** variant replicates the educational attainment of immigrants aged 25-44 in Canada arrived between 2011 and 2016, where about two third of immigrants are economic migrants selected according to their human capital (picked from the Canada 2016 Census).
- The **low education** variant, on its side, looks what would happen if immigrants' educational attainment is, at the EU-level, comparable to what observed in Italy in recent years, where immigrants are particularly low educated compared to other European countries (estimated from the Labour Force Survey 2011-2016).

Table S2: Educational attainment at age 25-44

Scenario	**Education**		
	Low	**Medium**	**High**
High	7%	20%	73%
Baseline	40%	27%	34%
Low	53%	35%	12%

Language spoken at home and religious affiliation

CEPAM-Mic also projects two ethno-cultural characteristics whose future evolution is highly dependent on immigration: religion and language spoken at home. At birth, religious affiliation and language spoken at home are taken directly from the mother, and are subsequently allowed to change during the life course. Transition rates for religious affiliation[5] are taken directly from the PEW projections on religion (9). Life course transition rates for language spoken at home are based on model schedules (22) calibrated using data from the ESS.

Labor force participation

The labor force participation module is described in detail in Marois et al. (5). The module is applied to individuals aged between 15 and 74. When a change occurs to the characteristic of an individual (age, education, duration of stay, etc.), the module determines probabilistically whether or not he/she participates in the labor force. The labor force participation status is imputed through a Monte-Carlo experiment in which a random number is compared to the probability of being active: a successful trial means that the simulated individual is active. Parameters are estimated from sex- and country-specific logit regressions on a binomial variable representing participation in the labor force, using pooled data from the 2010 to 2015 files of the annual EU Labour Force Survey. Equation 2 below describes the modeling of labor force participation (P):

Eq.S6 $$\text{logit}(P) = \beta_0 + \beta_1 AGE + \beta_2 EDU + \beta_3 YEAR + \beta_4(AGE * EDU) + \beta_5(AGE * YEAR) + \beta_6(EDU * YEAR) + \beta_7(AGE * EDU * YEAR) + \beta_8 IMMIG + \beta_9(IM15 * EDU)$$

[5] We consider the following four categories based on self-reported religious affiliation: Christians, Muslims, Unaffiliated and Other religion.

Where:

- $\beta_0 + \beta_1 + \beta_2 + \beta_4$ capture the joint effect of age and education on labor force participation rates[6]. Education is divided into 3 categories:
  1. Low (L): Lower secondary or less (ISCED 0, 1 ,and 2);
  2. Medium (M): Upper secondary completed (ISCED 3);
  3. High (H): Postsecondary (ISCED 97: 4, 5A, 5B and 6; ISCED 2011: 4, 5, 6, 7, 8);
- $\beta_3 + \beta_5 + \beta_6 + \beta_7$ capture the age and education specific trends in labor force participation;
- β_8 is a set of parameters for an immigration variable (IMMIG) combining place of birth[7], age at arrival and duration of stay. The variable is divided in five categories:
  1. Born in EU28;
  2. Born outside EU28, arrived before the age of 15;
  3. Born outside EU28, arrived after the age of 15, duration of stay < 5;
  4. Born outside EU28, arrived after the age of 15, 5 <= duration of stay < 10;
  5. Born outside EU28, arrived after the age of 15, 10 <= duration of stay;
- β_9 is a set of parameters estimating the labor force returns on education for migrants born outside the European Union and who arrived at the age of 15 or above (IM15).

Using β_0 to β_7, we estimated probabilities by sex, education and country for 2015 with those of the age group X-5 in 2010 (net from the immigration variable), and we computed entry and exit rates to build a labor force participation table for a synthetic cohort. The table is then used to assume net future participation rates, which show a notable increase in the participation rates of the population aged 50 to 74, in particular for women. In this paper, two assumptions are used concerning future trends in participation rates:

1. In the **baseline assumption**, entry and exit rates estimated with the method described above are kept constant throughout the projection. By cohort effect, resulting labor force participation rates increases in most countries, in particular for older women.
2. In the **Swedish assumption**, entry and exit rates are kept constant throughout the projection only for Sweden. For other countries, we assume that rates by age, sex, education reach those for the same group projected in Sweden in 2050. Values for intermediate years are interpolated following a logit curve.

In addition to the sex differential, regression models also account for another important source of inequalities in regards with labor force participation, as they explicitly take into account differentials between EU28 natives and foreign-born as well as the integration process through parameters for the duration of stay (β_8). The table S3 below shows the average value of β_8 for all EU countries. For men, labor force participation for immigrants born outside the EU is much lower than for EU-born individuals (β_8 =-0.936 for recent immigrants), although it improves with the number of years spent in the host country. After 10 years, the labor force participation rates of immigrants are close to the rates of EU-born individuals (β_8 = -0.223), and this holds true in most high immigration countries.

[6] The LFS does not provide information on labor force participation rates in the UK for the age group 70-74. It was assumed to be half of the value observed for the age group 65-69 for each education level.
[7] For Germany, the question on the country of birth is not asked in the LFS. We use the nationality as a proxy to distinguish EU28 migrants from international immigrants.

Table S3: Average value* of β_8 (IMMIG, see equation 2), EU28, 2010-2015		
	Women	**Men**
Born in EU28	Ref	Ref
Born outside EU28, arrived before the age of 15;	-0.220	-0.180
Born outside EU28, arrived after the age of 15, duration of stay <5;	-1.642	-0.936
Born outside EU28, arrived after the age of 15, 5<=duration of stay <10;	-1.258	-0.520
Born outside EU28, arrived after the age of 15, 10<=duration of stay;	-0.755	-0.223
*Values represent the average of parameters for the 28 countries. The level of significance varies among them, though most are highly significant (p<0.0001). Source: Marois, Sabourin, and Bélanger (5)		

Participation rates of immigrant women follow similar patterns as for immigrant men, although the gap compared to natives is wider. For recent immigrants born outside the European Union, the average value of β_8 for women is more than fifty percent lower than it is for men (β_8=-1.642, see Table S3). Even after 10 years, their participation rates are far below the rates of natives (β_8=-0.755). In fact, labor force participation rates of immigrant women do not reach the level of natives in any of the high immigration countries. Gender inequity in terms of labor force participation appears to be an issue affecting immigrant women more strongly than natives, which supports the double disadvantage theory (23, 24). These results are also consistent with evidences gathered in the U.S. by Antecol (25) and in Europe by Pessin & Arpino (26) concerning the role of cultural background in labor force integration: for some source regions, persistent gender gaps resist explanation based on socio-economic or institutional factors alone.

Large regional variations are, however, observed. In Denmark, where immigration inflows are dominated by asylum seekers and family reunification immigrants (27), the situation even appears to deteriorate with time for male immigrants, as β_8 drops from -1.078 for recent immigrants to -1.692 admitted more than ten years ago. In contrast, Spain emerges as a special case, with no clear differences in LFPRs between male immigrants and native-born population / men.

In addition, interaction parameters between the place of birth and education (β_9) allows the model to account for the fact that the return for education differs for immigrants and natives. Table S4 shows the average values of parameters for education alone (β_2) and for the interaction of education with a dichotomous variable for generation 1 migrants born outside the EU (β_9). Unsurprisingly, education is a major driver of labor force participation and has similar effects for both males and females. However, the effect of education is smaller for immigrants, especially for women. The positive effect on participation rates of having a high education level is about 50% lower for immigrant women when compared to native women. The impact of education on labor force participation is also reduced for male immigrants but less. These lower returns on education could be partly explained by lower quality degrees in source countries, as well as by cultural differences in the definition of gender roles (25, 28).

Table S4: Average value* of the parameters for education and its interaction with immigration, EU28, 2010-2015				
	Women		Men	
Level	EDU (β_2)	EDU*IM15 (β_9)	EDU (β_2)	EDU*IM15 (β_9)
Low	-1.753	0.817	-2.005	0.590
Medium	-0.753	0.467	-0.751	0.153

*Values represent the average of parameters for the 28 countries. The level of significance varies among them, though most are highly significant ($p<0.0001$).
Source: Marois, Sabourin, and Bélanger (5)

Taking these parameters into account allows the creation of assumptions on future participation rates with immigration differentials. In this paper, three variants are built the labor force participation of immigrants.

1. The **baseline integration** assumes continuation throughout the projection of β_8 and β_9. In other words, this assumption supposes no change in the integration of immigrants in labor force compared to what observed in recent years.
2. The **high integration** variant assumes that β_8 and β_9 converge to 0 for all countries by 2050. This assumption thus progressively removes the disadvantage of immigrants in the labor force. By 2050, for a same country, age, gender, and level of education immigrants and natives would have the same labor force participation rates.
3. The **low integration** variant, on its side, assumes a deterioration of the labor force participation of immigrants. β_8 and β_9 are assumed to converge by 2050 for all countries to those of Denmark in 2010-2015, which is the country in the EU where those parameters are the lowest (see parameters in table S5 and table S6).

Table S5: Value of β_8 for Denmark, used as convergence point in 2050 for the low integration assumption		
	Women	Men
Born in EU28	Ref	Ref
Born outside EU28, arrived before the age of 15;	-0.272	-0.301
Born outside EU28, arrived after the age of 15, duration of stay <5;	-1.078	-1.638
Born outside EU28, arrived after the age of 15, 5<=duration of stay <10;	-1.421	-1.236
Born outside EU28, arrived after the age of 15, 10<=duration of stay;	-1.692	-1.422

Source: Marois, Sabourin, and Bélanger (5)

Table S6: Value of β_9 for Denmark, used as convergence point in 2050 for the low integration assumption		
Education level	**Women**	**Men**
Low	0.681	0.865
Medium	0.314	0.235
Source: Marois, Sabourin, and Bélanger (5)		

Definition of scenarios

Immigration policies can try to impact the future by playing with three dimensions: the volume of immigration, the socioeconomic composition of migrants, or the extent to the integration of newcomers. This paper evaluates the effectiveness of these different options using 'what if' scenarios. All scenarios share the same assumptions (in terms of group-specific parameters) for fertility, mortality, domestic migration, education, language and religion shifts. However, because there are immigrant-specific considerations for some of the 'events' occurring during the microsimulation such as fertility and because immigrants differ from natives in terms of sociodemographic composition, the total fertility rate or total life expectancy varies between scenarios, despite using same parameters for sub-groups. The table S7 summarizes the scenarios.

Table S7: Summary of scenarios

Component	i. Baseline	ii. Baseline/ Swedish_LF	iii. Canadian	iv. Canadian/ Swedish_LF	v. Canadian/ Hi_Int	vi. Canadian/ Lo_Ed/Lo_Int	vii. Japenese
Volume of immigration	10M/5 years	10M/5 years	20M/5 years	20M/5 years	20M/5 years	20M/5 years	1.2M/5 years
Educational composition of immigrants	Same as recent immigrants	Same as recent immigrants	Same as recent immigrants in Canada	Same as recent immigrants in Canada	Same as recent immigrants in Canada	Same as recent immigrants in Italy	Same as recent immigrants in Canada
Integration of immigrants	Average of 2010-2015	Average of 2010-2015	Average of 2010-2015	Average of 2010-2015	Rates reach those of EU born	Rates reach those of immigrants in Denmark	Average of 2010-2015
Labor force participation trends	Constant entry and exit rates	Rates reach those of Sweden	Constant entry and exit rates	Rates reach those of Sweden	Constant entry and exit rates	Constant entry and exit rates	Constant entry and exit rates
Fertility	Slight increase in the TFR from 1.6 to 1.8						
Mortality	Continuous improvement in life expectancy						
Educational attainment	Past trends continue, constant parameters for sociocultural characteristics						
Domestic migration	Average of 2013-2016						
Language and religion shift	Baseline rates						

The (i) **baseline** scenario is "business as usual". Parameters for immigration, integration, and labor force participation follow recent trends. The (ii) **Baseline/Swedish_LF** is the scenario supposing continued increases in labor force participation, in particular among women and elderly, to the levels observed in Sweden today. Immigration-related parameters are the same as scenario 1. The (iii) **Canadian** scenario supposes that the EU adopts an immigration system similar to the Canadian one in which immigrants are numerous and selected according to their human capital. Immigration volumes are thus high and immigrants are more educated. Since immigrants in Canada still face economic integration issues despite its selection systems (29), the integration assumption in this scenario is the same as in the baseline (though overall labor force participation rates for immigrants are higher considering their more favorable educational attainment). The (iv) **Canadian/Swedish_LF** scenario combines both an immigration system similar to Canada's and efficient policies to increase labor force participation of the population. The (v) **Canadian/Hi_Int** scenario assumes also an immigration system similar to Canada's, but with more success with the economic integration of immigrants. The (vi) **Canadian/Lo_Ed/Lo_Int** scenario shows what would happen if the EU has high level of immigrants, but fails in their selection and integration. Finally, the (vii) **Japanese** scenario shows what would happen if the EU adopts more selective migration policies to both emphasize skills and substantially reduce overall volume of flows.

Appendix 2 – Detailed results by scenarios and country, 2015, 2030 and 2060

Age-dependency ratio (ADR)

		2030							2060						
Country	2015	Baseline	Baseline/ Swedish_LF	Canadian	Canadian/ Swedish_LF	Canadian/ Hi_Int	Canadian/ Lo_Ed/Lo_Int	Japanese	Baseline	Baseline/ Swedish_LF	Canadian	Canadian/ Swedish_LF	Canadian/ Hi_Int	Canadian/ Lo_Ed/Lo_Int	Japanese
Austria	0.49	0.62	0.62	0.61	0.61	0.61	0.61	0.64	0.82	0.82	0.74	0.74	0.74	0.74	0.97
Belgium	0.54	0.65	0.65	0.64	0.64	0.64	0.64	0.66	0.80	0.80	0.74	0.74	0.73	0.74	0.91
Bulgaria	0.51	0.62	0.62	0.61	0.61	0.61	0.61	0.63	0.81	0.81	0.77	0.77	0.77	0.76	0.93
Cyprus	0.45	0.58	0.58	0.56	0.56	0.57	0.56	0.58	0.88	0.88	0.79	0.79	0.78	0.77	1.20
Czech Republic	0.49	0.60	0.60	0.60	0.60	0.59	0.60	0.60	0.92	0.92	0.86	0.86	0.85	0.85	1.00
Germany	0.52	0.69	0.69	0.67	0.67	0.67	0.67	0.71	0.83	0.83	0.75	0.75	0.74	0.74	0.99
Denmark	0.55	0.66	0.66	0.65	0.65	0.66	0.65	0.68	0.75	0.75	0.69	0.69	0.69	0.70	0.87
Estonia	0.53	0.63	0.63	0.62	0.62	0.64	0.62	0.63	0.83	0.83	0.78	0.78	0.75	0.77	0.93
Spain	0.51	0.61	0.61	0.59	0.59	0.59	0.59	0.63	1.04	1.04	0.89	0.89	0.89	0.90	1.29
Finland	0.57	0.73	0.73	0.72	0.72	0.71	0.72	0.74	0.81	0.81	0.75	0.75	0.76	0.75	0.88
France	0.59	0.71	0.71	0.70	0.70	0.70	0.70	0.72	0.82	0.82	0.77	0.77	0.76	0.76	0.89
Greece	0.55	0.67	0.67	0.66	0.66	0.66	0.66	0.67	1.09	1.09	0.98	0.98	0.99	1.00	1.24
Croatia	0.50	0.64	0.64	0.64	0.64	0.63	0.64	0.65	0.91	0.91	0.86	0.86	0.87	0.87	1.03
Hungary	0.48	0.56	0.56	0.55	0.55	0.56	0.56	0.57	0.79	0.79	0.73	0.73	0.73	0.74	0.88
Ireland	0.52	0.58	0.58	0.58	0.58	0.58	0.60	0.58	0.84	0.84	0.74	0.74	0.75	0.74	1.00
Italy	0.55	0.67	0.67	0.65	0.65	0.66	0.65	0.68	0.91	0.91	0.82	0.82	0.82	0.82	1.06
Lithuania	0.50	0.70	0.70	0.70	0.70	0.70	0.70	0.73	0.93	0.93	0.83	0.83	0.82	0.81	1.04
Luxemburg	0.45	0.55	0.55	0.58	0.58	0.58	0.58	0.59	0.79	0.79	0.69	0.69	0.71	0.67	0.80
Latvia	0.52	0.64	0.64	0.64	0.64	0.62	0.63	0.64	0.88	0.88	0.83	0.83	0.82	0.83	1.03
Malta	0.48	0.59	0.59	0.54	0.54	0.55	0.55	0.61	0.69	0.69	0.63	0.63	0.61	0.60	0.84
Netherlands	0.53	0.67	0.67	0.66	0.66	0.66	0.66	0.68	0.80	0.80	0.73	0.73	0.72	0.73	0.91
Poland	0.44	0.61	0.61	0.61	0.61	0.60	0.62	0.62	0.89	0.89	0.81	0.81	0.81	0.82	0.98
Portugal	0.53	0.66	0.66	0.66	0.66	0.66	0.66	0.67	1.01	1.01	0.96	0.96	0.96	0.95	1.10
Romania	0.48	0.56	0.56	0.56	0.56	0.56	0.56	0.56	0.84	0.84	0.78	0.78	0.78	0.77	0.93
Sweden	0.59	0.66	0.66	0.65	0.65	0.64	0.63	0.67	0.75	0.75	0.69	0.69	0.68	0.68	0.90
Slovenia	0.49	0.67	0.67	0.67	0.67	0.67	0.66	0.69	0.92	0.92	0.82	0.82	0.82	0.80	1.13
Slovakia	0.41	0.56	0.56	0.56	0.56	0.56	0.56	0.56	0.88	0.88	0.87	0.87	0.87	0.88	0.89
United Kingdom	0.55	0.64	0.64	0.63	0.63	0.63	0.64	0.65	0.77	0.77	0.71	0.71	0.71	0.72	0.86
European Union	**0.53**	**0.65**	**0.65**	**0.64**	**0.64**	**0.64**	**0.64**	**0.66**	**0.85**	**0.85**	**0.77**	**0.77**	**0.77**	**0.77**	**0.97**

Labor force dependency ratio (LFDR)

Country	2015	2030							2060						
		Baseline	Baseline/ Swedish_LF	Canadian	Canadian/ Swedish_LF	Canadian/ Hi_Int	Canadian/ Lo_Ed/Lo_Int	Japanese	Baseline	Baseline/ Swedish_LF	Canadian	Canadian/ Swedish_LF	Canadian/ Hi_Int	Canadian/ Lo_Ed/Lo_Int	Japanese
Austria	0.97	1.11	1.00	1.09	0.99	1.05	1.13	1.11	1.32	1.04	1.21	0.96	1.11	1.33	1.41
Belgium	1.28	1.42	1.18	1.40	1.17	1.34	1.43	1.38	1.58	1.06	1.49	1.00	1.33	1.64	1.62
Bulgaria	1.17	1.25	1.06	1.25	1.06	1.24	1.30	1.23	1.47	1.00	1.45	0.98	1.38	1.62	1.57
Cyprus	0.92	1.04	0.92	0.99	0.89	0.96	1.04	1.01	1.38	1.07	1.33	1.01	1.16	1.58	1.64
Czech Republic	0.98	1.03	0.95	1.03	0.95	1.01	1.06	1.02	1.36	1.09	1.29	1.03	1.27	1.46	1.43
Germany	0.94	1.10	1.03	1.08	1.02	1.03	1.12	1.07	1.30	1.08	1.21	1.01	1.07	1.33	1.34
Denmark	0.93	1.07	1.01	1.05	1.00	1.00	1.11	1.04	1.21	1.06	1.16	1.02	0.98	1.25	1.20
Estonia	0.95	0.99	0.93	0.99	0.93	1.00	1.02	0.99	1.16	1.01	1.11	0.96	1.03	1.29	1.25
Spain	1.01	1.12	1.01	1.08	0.97	1.07	1.13	1.15	1.61	1.23	1.40	1.06	1.42	1.71	1.93
Finland	1.02	1.19	1.08	1.16	1.06	1.13	1.19	1.17	1.26	1.02	1.18	0.96	1.14	1.34	1.28
France	1.22	1.33	1.17	1.31	1.17	1.28	1.34	1.31	1.38	1.03	1.34	1.01	1.26	1.45	1.43
Greece	1.23	1.47	1.17	1.43	1.15	1.43	1.48	1.48	1.99	1.26	1.84	1.14	1.87	2.08	2.25
Croatia	1.25	1.26	1.09	1.24	1.07	1.23	1.30	1.28	1.57	1.08	1.49	1.07	1.44	1.63	1.67
Hungary	1.17	1.06	0.98	1.05	0.96	1.04	1.07	1.06	1.33	0.96	1.26	0.90	1.23	1.36	1.42
Ireland	1.15	1.17	1.01	1.19	1.02	1.15	1.32	1.14	1.47	1.08	1.38	1.01	1.25	1.69	1.57
Italy	1.38	1.29	1.17	1.25	1.14	1.25	1.32	1.31	1.38	1.10	1.23	1.00	1.23	1.52	1.50
Lithuania	1.00	1.10	1.07	1.11	1.09	1.09	1.16	1.14	1.33	1.11	1.25	1.05	1.22	1.43	1.45
Luxemburg	1.02	1.02	0.98	0.99	1.03	0.99	1.02	1.03	1.24	1.00	1.10	0.86	1.16	1.15	1.23
Latvia	0.99	1.04	1.00	1.05	1.01	0.99	1.06	1.01	1.24	1.06	1.19	1.03	1.11	1.33	1.32
Malta	1.17	0.95	1.01	0.86	0.95	0.84	0.92	0.94	0.98	0.94	0.89	0.91	0.76	0.96	1.02
Netherlands	0.90	1.00	1.00	0.99	0.99	0.93	1.03	0.97	1.14	1.09	1.06	1.02	0.92	1.12	1.13
Poland	1.10	1.11	1.01	1.10	1.01	1.09	1.15	1.12	1.38	1.06	1.27	0.99	1.26	1.51	1.45
Portugal	1.00	1.05	1.03	1.04	1.03	1.03	1.05	1.07	1.38	1.22	1.34	1.17	1.32	1.41	1.48
Romania	1.16	1.15	1.01	1.15	1.02	1.14	1.17	1.15	1.40	1.03	1.36	1.00	1.26	1.42	1.47
Sweden	0.90	0.89	0.94	0.88	0.94	0.84	0.89	0.87	0.96	1.02	0.89	0.96	0.82	0.96	1.03
Slovenia	1.07	1.25	1.15	1.22	1.13	1.21	1.24	1.29	1.51	1.15	1.37	1.07	1.33	1.51	1.77
Slovakia	0.99	0.97	0.93	0.98	0.94	0.97	0.98	0.97	1.34	1.04	1.32	1.03	1.33	1.35	1.36
United Kingdom	0.98	1.05	0.98	1.04	0.97	1.01	1.10	1.05	1.16	1.00	1.11	0.95	1.02	1.25	1.20
European Union	**1.08**	**1.15**	**1.05**	**1.13**	**1.04**	**1.10**	**1.17**	**1.14**	**1.33**	**1.07**	**1.24**	**1.00**	**1.16**	**1.40**	**1.41**

Productivity-weighted labor force dependency ratio (PWLFDR)

Country	2015	2030							2060						
		Baseline	Baseline/ Swedish_LF	Canadian	Canadian/ Swedish_LF	Canadian/ Hi_Int	Canadian/ Lo_Ed/Lo_Int	Japanese	Baseline	Baseline/ Swedish_LF	Canadian	Canadian/ Swedish_LF	Canadian/ Hi_Int	Canadian/ Lo_Ed/Lo_Int	Japanese
Austria	0.90	1.00	0.91	0.95	0.87	0.91	1.04	0.98	1.14	0.91	0.99	0.79	0.90	1.19	1.15
Belgium	1.17	1.22	1.03	1.16	0.99	1.12	1.26	1.18	1.26	0.88	1.15	0.80	1.03	1.36	1.25
Bulgaria	1.07	1.12	0.97	1.12	0.96	1.11	1.19	1.11	1.31	0.91	1.27	0.87	1.19	1.50	1.39
Cyprus	0.80	0.86	0.77	0.80	0.73	0.77	0.90	0.83	1.15	0.90	1.03	0.79	0.89	1.37	1.26
Czech Republic	0.93	0.96	0.89	0.94	0.88	0.92	1.00	0.94	1.27	1.05	1.15	0.96	1.13	1.39	1.32
Germany	0.84	0.96	0.90	0.92	0.87	0.87	1.00	0.92	1.11	0.94	0.98	0.82	0.86	1.18	1.08
Denmark	0.87	0.97	0.92	0.92	0.88	0.87	1.04	0.93	1.08	0.95	0.98	0.86	0.82	1.18	1.03
Estonia	0.79	0.83	0.80	0.82	0.78	0.83	0.88	0.83	0.96	0.86	0.90	0.80	0.85	1.13	1.02
Spain	0.97	1.01	0.90	0.93	0.84	0.92	1.03	1.02	1.32	1.03	1.07	0.83	1.08	1.48	1.47
Finland	0.89	1.03	0.95	0.99	0.91	0.96	1.06	1.01	1.08	0.89	0.97	0.80	0.94	1.17	1.05
France	1.10	1.13	1.01	1.09	0.99	1.06	1.16	1.10	1.10	0.84	1.04	0.80	0.98	1.20	1.10
Greece	1.12	1.27	1.01	1.21	0.97	1.21	1.29	1.26	1.59	1.03	1.41	0.90	1.44	1.70	1.70
Croatia	1.20	1.16	1.01	1.11	0.97	1.11	1.21	1.17	1.38	0.96	1.26	0.92	1.21	1.47	1.43
Hungary	1.07	0.95	0.88	0.93	0.86	0.93	0.98	0.96	1.19	0.86	1.09	0.78	1.06	1.28	1.25
Ireland	0.94	0.93	0.81	0.94	0.82	0.91	1.11	0.92	1.14	0.86	1.07	0.80	0.96	1.44	1.21
Italy	1.44	1.24	1.14	1.15	1.07	1.15	1.28	1.24	1.18	0.97	0.99	0.83	0.98	1.31	1.22
Lithuania	0.77	0.88	0.86	0.88	0.86	0.86	0.95	0.91	1.07	0.91	1.00	0.86	0.97	1.25	1.17
Luxemburg	0.94	0.89	0.86	0.86	0.89	0.87	0.92	0.91	1.02	0.84	0.89	0.71	0.93	1.00	0.99
Latvia	0.90	0.92	0.89	0.92	0.89	0.87	0.96	0.90	1.05	0.90	1.00	0.88	0.92	1.19	1.10
Malta	1.19	0.89	0.95	0.75	0.83	0.72	0.90	0.84	0.82	0.81	0.69	0.71	0.59	0.85	0.80
Netherlands	0.84	0.88	0.88	0.84	0.85	0.80	0.92	0.84	0.94	0.90	0.85	0.81	0.73	0.97	0.90
Poland	0.98	0.96	0.89	0.94	0.87	0.92	1.01	0.96	1.17	0.93	1.05	0.84	1.03	1.32	1.20
Portugal	1.11	1.03	1.02	1.01	1.00	1.00	1.04	1.05	1.15	1.02	1.09	0.96	1.08	1.19	1.21
Romania	1.14	1.07	0.96	1.07	0.96	1.06	1.11	1.07	1.22	0.91	1.16	0.86	1.08	1.28	1.29
Sweden	0.80	0.78	0.83	0.74	0.79	0.71	0.82	0.75	0.83	0.89	0.73	0.79	0.67	0.89	0.85
Slovenia	1.02	1.15	1.06	1.08	1.01	1.06	1.17	1.16	1.37	1.06	1.16	0.91	1.12	1.42	1.49
Slovakia	0.92	0.89	0.86	0.90	0.86	0.89	0.90	0.89	1.22	0.96	1.20	0.94	1.21	1.24	1.24
United Kingdom	0.89	0.91	0.85	0.89	0.83	0.86	0.99	0.91	0.96	0.83	0.90	0.78	0.83	1.11	0.98
European Union	**1.00**	**1.02**	**0.94**	**0.98**	**0.90**	**0.95**	**1.06**	1.01	**1.11**	**0.91**	**1.00**	**0.82**	**0.93**	**1.22**	**1.14**

References

1. A. Bélanger, P. Sabourin, *Microsimulation and Population Dynamics. An Introduction to Modgen 12* (Springer International Publishing, 2017).

2. A. Bélanger, P. Sabourin, G. Marois, J. Van Hook, S. Vézina, A framework for the prospective analysis of ethno-cultural super-diversity. *Demographic Research* **41**, 293–330 (2019).

3. E. Van Imhoff, W. Post, Microsimulation methods for population projection. *Popul* **10**, 97–138 (1998).

4. M. Spielauer, What is Social Science Microsimulation? *Social Science Computer Review* **29**, 9–20 (2010).

5. G. Marois, P. Sabourin, A. Bélanger, Implementing Dynamics of Immigration Integration in Labor Force Participation Projection in EU28. *Population Research and Policy Review* (2019) https:/doi.org/10.1007/s11113-019-09537-y.

6. G. Marois, P. Sabourin, A. Bélanger, Forecasting Human Capital of EU Member Countries Accounting for Sociocultural Determinants. *Journal of Demographic Economics* **85**, 231–269 (2019).

7. M. Potančoková, G. Marois, Projecting the future births in the EU28 with fertility differentials reflecting women's educational and migrant characteristics. *Vienna Yearbook of Population Research* (Forthcoming).

8. S. van Buuren, K. Groothuis-Oudshoorn, mice: Multivariate Imputation by Chained Equations in R. *Journal of Statistical Software; Vol 1, Issue 3 (2011)* (2011).

9. C. Hackett, P. Connor, M. Stonawski, V. Skirbekk, "The Future of World Religions: Population Growth Projections, 2010-2050. Why Muslims are Rising Fastest and the Unafiliated Are Shrinking as a Share of the World's Population" (PEW Research Center, 2015).

10. WIC, Wittgenstein Centre Data Explorer Version 1.2 (2015) (March 18, 2015).

11. W. Lutz, A. V. Goujon, S. KC, M. Stonawski, N. Stilianakis, "Demographic and human capital scenarios for the 21st century: 2018 assessment for 201 countries" (Publications Office of the European Union, 2018) https:/doi.org/10.2760/41776.

12. G. Caselli, S. Drefahl, C. Wegner-Siegmundt, M. Luy, "Future mortality in low mortality countries" in *World Population and Human Capital in the 21st Century*, W. Lutz, W. P. Butz, S. KC, Eds. (Oxford University Press, 2014), pp. 226–272.

13. S. Basten, T. Sobotka, K. Zeman, "Future fertility in low fertility countries" in *World Population and Human Capital in the 21st Century*, W. Lutz, W. P. Butz, S. KC, Eds. (Oxford University Press, 2014), pp. 39–146.

14. A. Heath, Y. Brinbaum, Guest editorial: Explaining ethnic inequalities in educational attainment. *Ethnicities* **7**, 291–304 (2007).

15. C. Hirschman, The educational enrollment of immigrant youth: a test of the segmented-assimilation hypothesis. *Demography* **38**, 317–336 (2001).

16. R. T. Riphahn, Cohort effects in the educational attainment of second generation immigrants in Germany: An analysis of census data. *Journal of Population Economics* **16**, 711–737 (2003).

17. Y. Shavit, M. Yaish, E. Bar-haim, "The persistence of persistent inequality" in *From Origin to Destination: Trends and Mechanisms in Social Stratification Research*, (University of Chicago Press., 2007), pp. 42–51.

18. W. Lutz, W. P. Butz, S. KC, Eds., *World Population and Human Capital in the Twenty-First Century* (Oxford University Press, 2014).

19. P. Rees, P. Wohland, P. Norman, N. Lomax, "Sub-national projection methods for Scotland and Scottish areas: A review and recommendations" (White Rose University Consortium, University of Leeds, 2015).

20. A. Rogers, L. J. Castro, Model migration schedules. *IIASA Research Report* **81**, 1–160 (1981).

21. J. Raymer, A. Wiśniowski, J. J. Forster, P. W. F. Smith, J. Bijak, Integrated Modeling of European Migration. *Journal of the American Statistical Association* **108**, 801–819 (2013).

22. P. Sabourin, A. Bélanger, The Dynamics of Language Shift in Canada. *Population* **Vol. 70**, 727–757 (2015).

23. M. Boyd, At a disadvantage: the occupational attainments of foreign born women in Canada. *Int Migr Rev* **18**, 1091–1119 (1984).

24. G. Ballarino, N. Panichella, The occupational integration of migrant women in Western European labour markets. *Acta Sociologica*, 000169931772344 (2017).

25. H. Antecol, An examination of cross-country differences in the gender gap in labor force participation rates. *Labour Economics* **7**, 409–426 (2000).

26. L. Pessin, B. Arpino, Navigating between two cultures: Immigrants' gender attitudes toward working women. *Demographic Research* **38**, 967–1016 (2018).

27. T. Liebig, "The Labour Market Integration of Immigrants in Denmark" (Organisation for Economic Co-operation and Development, 2007) https:/doi.org/10.1787/233783261534 (September 8, 2017).

28. R. Inglehart, P. Norris, *Rising Tide: Gender Equality and Cultural Change Around the World* (Cambridge University Press, 2003).

29. A. Aydemir, Immigrant selection and short-term labor market outcomes by visa category. *Journal of Population Economics* **24**, 451–475 (2011).