WP 4.2
16 October 2013

UNITED NATIONS STATISTICAL COMMISSION and ECONOMIC COMMISSION FOR EUROPE

STATISTICAL OFFICE OF THE EUROPEAN UNION (EUROSTAT)

Joint Eurostat/UNECE Work Session on Demographic Projections
organised in cooperation with Istat
(29-31 October 2013, Rome, Italy)

Item 4 – Assumptions on future migration

Introducing duration dependant emigration in the DREAM population projection model

Marianne Frank Hansen, Danish Rational Economic Agents Model (DREAM)

Abstract

During the recent decade, changes in immigration flows and immigration behaviour are important sources to explain changes in the composition of the resident immigrant population with respect to duration of residence. Considering that demographic behaviour varies considerably with the length of duration, this challenges the baseline assumption of not considering duration of residence when determining future demographic flows. This working paper explains the consequences of allowing forecasted emigration of immigrants, i.e. re-emigration, to depend on duration of residence and investigates whether including this characteristic enhances projection accuracy when facing shifts in immigration structure.

The propensity to re-emigrate decreases with duration of residence. Typically emigration probabilities for individuals having immigrated within the last two years lie above the average re-emigration probability, whereas the propensity to re-emigrate lies below average when duration of residence exceeds two years. Using constant emigration probabilities depending on gender, age, origin, and number of years of duration is shown to lead to an increase in the immigrant population compared to the baseline scenario omitting emigration by duration of residence.

The implications with respect to projection accuracy are assessed by performing sequential within-data population projections, respectively involving and omitting duration dependant re-emigration. Finding that future shifts in immigration behaviour severely challenge projection accuracy when taking duration of residence into account, it is suggested that duration dependant emigration should be applied with caution.

DREAM, Danish Rational Economic Agents Model. Amaliegade 44, 1256 København K, e-mail: mfr@dreammodel.dk, web: www.dreammodel.dk

1. Introduction

For foreign nationals the possibilities of taking residence in Denmark are primarily affected by law and by the situation in the country of origin. The difficulties in projecting changes in these factors contribute significantly to the challenge of determining future immigration and future immigrant behaviour. During the past decade an increase is observed in immigration from especially Western countries. This is mainly due to legislation easing the access to the Danish labour market for citizens from Eastern European countries. Since annual immigration from non-Western countries is fairly constant during the same period, a change is induced in the composition of origin of resident immigrants. This is reflected in a shift in the pattern of residence permits, which are recently being granted primarily on the grounds of work or study rather than on the grounds of asylum and family reunification. Work- and study-warranted residence permits are generally associated with a shorter duration of residence than other permits, leading to a change in the composition of the immigrant population with respect to duration of residence.

Considering that demographic behaviour varies considerably with the length of duration, this challenges the DREAM baseline assumption of not considering duration of residence when determining future demographic flows. The following sections explain the consequences of allowing forecasted emigration of immigrants, i.e. re-emigration, to depend on duration of residence and investigates whether including this characteristic enhances projection accuracy when facing the aforementioned shifts in immigration structure. Taking the duration dimension into account results in missing observations for some combinations of gender, age, and origin. As a consequence implementation of the duration dimension is being limited to re-emigration taking just five years of duration of residence into account. Data mining techniques can potentially be applied to investigate the consequences of including more than five years of residency as well as the effects of duration dependency of various other types of demographic behaviour. However, this is reserved for further studies.

In general the propensity to re-emigrate decreases with duration of residence. Typically emigration probabilities for individuals having immigrated within the last two years lie above the average re-emigration probability, whereas the propensity to re-emigrate lies below average when duration of residence exceeds two years. Using constant emigration probabilities depending on gender, age, origin, and number of years of duration in a group-based population projection model is shown to lead to an increase in the immigrant population. This is due to the fact that the effect of lower than average emigration propensities for those with long residencies dominates because of composition effects. In addition, a change is observed in the age-composition of immigrants, hereby altering future descendant population.

By performing sequential within-data[1] population projections with a five-year forecast horizon, respectively involving and omitting duration dependant re-emigration, the challenges of including this demographic characteristic are assessed. Finding that a shift in immigration behaviour severely challenges projection accuracy when taking duration into account, it is suggested that duration dependant re-emigration should be used with caution.

[1] Projections are based on constant emigration probabilities using a three-year average prior to each of the years 1995-2007, respectively involving and omitting duration.

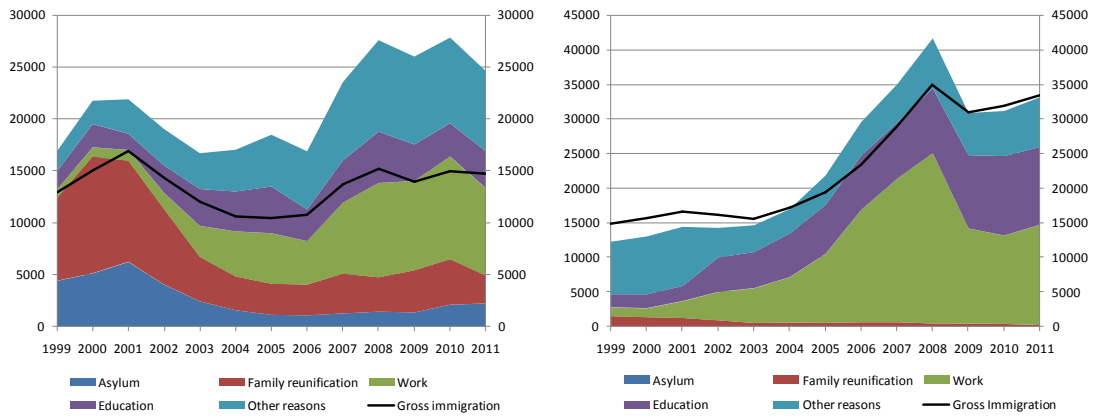
The paper is organized as follows. Section 2 is dedicated to a brief overview of the historical development in the immigration pattern and the hereby induced change in the composition of the immigrant population across origin and duration of residence. Further the change over time and the variation in emigration propensities across duration of residence are illustrated. Following a description of the baseline projection model in section 3, emigration by duration of residence is introduced, and the nature of data is described. The consequences of including duration dependant emigration behaviour in the demographic projection are outlined in section 4. By performing within-data projections, respectively involving and omitting duration dependant emigration, projection accuracy is assessed in section 5. The conclusion is presented in section 6.

2. An overview of immigration to Denmark in recent decades

Business cycle effects, legislation, and the political environment in the country of origin are among the most important factors determining not only the quantity of immigration but also the composition of residence permits. During the last two decades a significant change in the pattern of residence permits can be observed and alterations are typically easily identified with known changes in the aforementioned factors. From initially being granted on the grounds of asylum and family reunification, residence permits are now primarily associated with work or study related stays, cf. Figure 2.1. This shift is particularly evident for immigrants from non-Western countries since resident permits given on the grounds of asylum or family reunification are rare, when regarding Western immigration.

Figure 2.1. Residence permits by permit status and total gross immigration

a) Non-Western immigrants *b) Western immigrants*



Note: The figure contains all resident permits warranted within a year to foreign citizens. If an individual experiences a change in residence permit status, this will count as an additional permit in the data. Equivalently a single individual will count as two immigration events if entering, exiting and re-entering with the same year. Data on immigration include descendants of immigrants, re-immigrating and holding a foreign citizenship.

Source: Statistics Denmark StatBank table VAN6 and INDVAN.

For non-Western immigrants the development in residence permits granted on the grounds of asylum or family reunification is easily tracked to the civil war in the former Yugoslavian Republic in the mid-

nineties, conflicts within the African continent, and most recently the political situation in Iraq, Afghanistan, and Syria. By imposing further requirements that must be met in order to obtain a residence permit, changes in the Danish immigration legislation have induced a decrease in the number of permits being granted on the grounds of family reunification and asylum following the year of legislation change in 2002. From 2003 to 2009 legislation gradually easing the access to the Danish labour market for citizens from certain Eastern European countries[2] is introduced, leading to an increase in the number of work-warranted residence permits given to Western immigrants. This increase is likely enhanced by favourable business-cycle effects since also evident for immigrants from non-Western countries. Business-cycle changes are also likely to be partly responsible for the decrease in work-warranted permits observed after 2008.

The number of residence permits warranted within a certain year is unlikely to represent the number of immigrants arriving throughout the same year. This is partly explained by the fact that a residence permit is not necessarily warranted in the year of immigration and by the possibility of an individual experiencing a change in residence permit status. The latter will lead to at least two residence permits being associated with a single immigration event. However, the development in the number of residence permits being warranted will to a large extent mirror the development in immigration over time, cf. Figure 2.1. Typically the number of immigration events are inferior to the number of residence permits warranted.

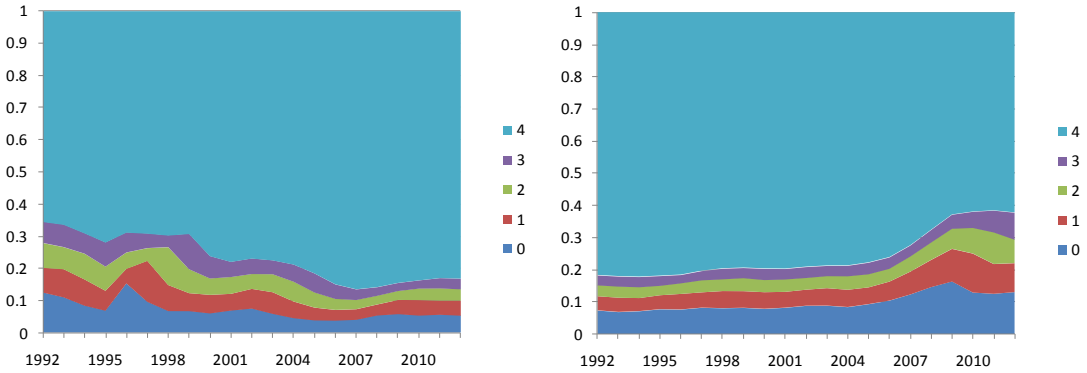
Changes in the pattern of residence permits and changes in the immigration quantity are significant sources to explain alterations in the distribution of duration of residence within the immigrant population. Work- and study-warranted residence permits are generally associated with a shorter duration of residence than other permits. A change in the grounds on which residence permits are being warranted is therefore likely to lead to changes in average immigrant behaviour and thus to changes in the distribution of the immigrant population on duration of residence. Immigrants having not previously been residing in Denmark are assigned a duration of residence of zero years on arrival. An increase in the annual flow of this type of immigration is therefore also likely to change the composition of the immigrant population with respect to duration of residence.

As can be seen from Figure 2.2, a change in the duration of residence distribution is evident within both the Western and the non-Western immigrant population during the recent decades. To a certain extent the development reflects the change in gross immigration depicted in Figure 2.1. Especially for Western immigrants, the recent persistent increase in annual immigration has contributed to an increasing share of short residencies.

[2] In order to restrict the access to the Danish labour market for citizens from the eight Eastern European countries joining the EU in May 2004, legislation was imposed in 2003 requiring obtainment of a working permit prior to commencing employment in Denmark. The restrictions were gradually remitted in the following years.

Figure 2.2. Immigrant population by origin and five categories of duration of residence

a) Non-Western immigrants *b) Western immigrants*



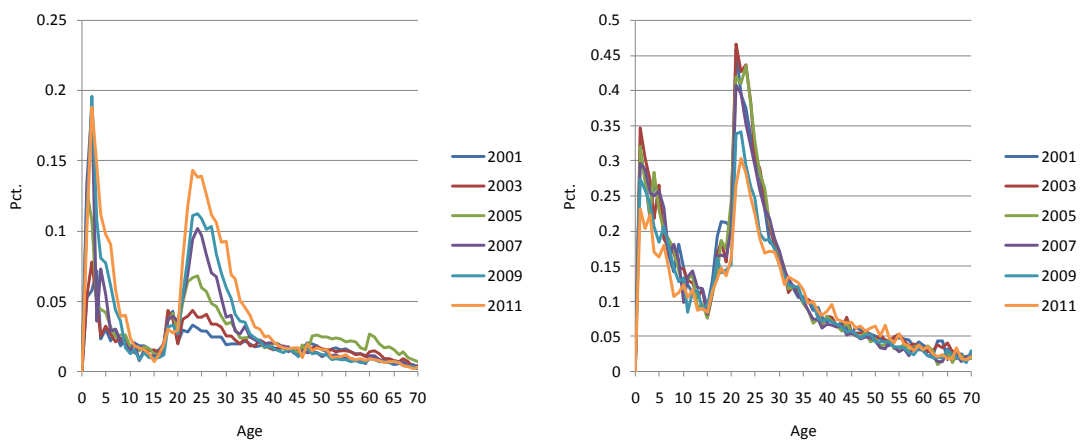
Note: The duration of residence refers to the status at the beginning of the year. Immigrants arriving during year t-1 are holding a duration of stay of zero years at the beginning of year t, indicating a duration of residence less than one year. The category indexed by a duration of residence of "4" comprises the immigrant population with a duration of stay of 4 years and above.

Source: Statistics Denmark.

As mentioned changes in demographic behaviour are also likely to be at least partly explained by the change in the pattern of residence permits. Figure 2.3 illustrates the change in emigration propensities during the recent decade for male immigrants from non-Western and Western countries. The emigration propensities for non-Western immigrants have been increasing over time, while on the contrary emigration propensities for Western immigrants have been decreasing. Consequently this has reduced the difference between non-Western and Western emigration propensities, the latter typically being by far the largest. Though omitted for brevity, this tendency is also evident for women.

Figure 2.3. Change over time in male emigration propensities by age and origin

a) Non-Western immigrants *b) Western immigrants*



Note: The depicted propensities of year t are calculated as a three-year average of the actual propensities from year t-2 to t.

Source: Statistics Denmark.

Due to registration issues, data describing each individual in the resident immigrant population by gender, age, and duration of residence, cf. Figure 2.2, cannot be combined with data describing types of residence permits. This disables immediate verification of the hypothesized relationship between types of residence permits and duration of residence. However, strong support for the alleged correlation is found in Statistics Denmark (2008, 2011). Based on a set of special assumptions these papers link duration of residence to residence permits granted on the grounds of fugitive status and non-fugitive status respectively[3]. This suggests that immigrants holding resident permits granted on the grounds of asylum have a long duration of residence[4].

As depicted in Figure 2.4 the emigration propensities for short residencies typically lie above the average propensity, whereas the propensity to emigration lies below average for long residencies. This pattern is valid for any given point in time, though the dispersion from the average might decrease or increase. According to Figure 2.1 the annual immigration flow from non-Western and Western countries have more or less been increasing since 2005, contributing to an increase in the share of the immigrant population holding a short duration of stay, cf. Figure 2.2. Assuming that the propensities to emigrate by duration of residence were constant during the same period and ranked according to the pattern in Figure 2.4, composition effects would lead to a long run increase in the average emigration propensity[5]. From Figure 2.3 this is only apparent for non-Western immigrants, thus suggesting that the change in the average emigration propensity is not only explained by increasing immigration and composition effects, but should also be attributed to a change in emigration behaviour by duration of residence. Expanding Figure 2.4 to comprise the development during the recent decade will in fact confirm that the tendency from Figure 2.3 is reproduced when regarding the development in emigration behaviour by duration of residence (not illustrated). I.e. the direction of movement of the average emigration propensities can be identified in emigration propensities representing both short and long residencies. Since the change in the pattern of residence permits not only contributes to a change in the average emigration propensity, but also induces changes in emigration by duration of residence, variation not explained by duration of stay exits throughout the historical period.

---

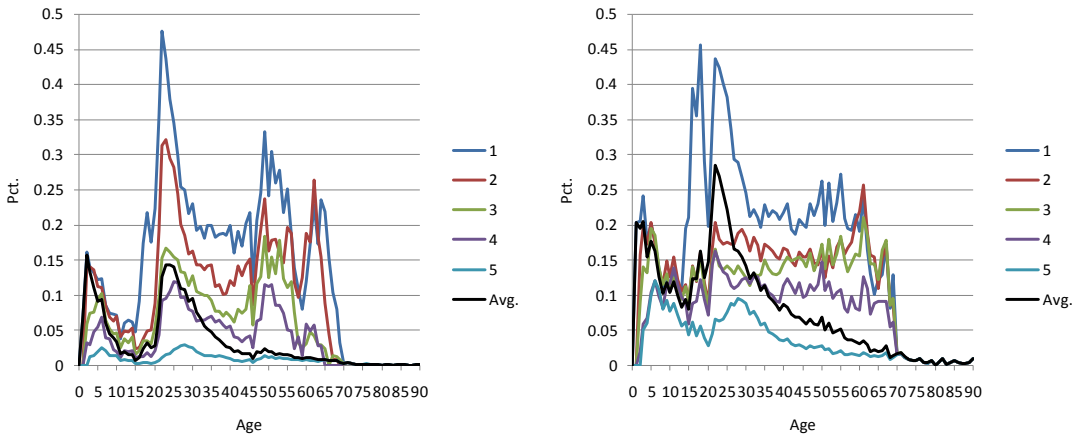
[3] Information regarding the type of resident permits granted to immigrants arriving prior to 1997 is typically not referred to a social security number. However, if the country of origin is classified as a refugee country in the year of immigration, the immigrant is classified as having obtained a residence permit on the grounds of fugitive status, cf. Statistics Denmark (2008).

[4] Statistics Denmark (2011): Section 1.11 pp. 41.

[5] Note that this is also contingent on the maximum residency category being the dominant in the duration distribution of the immigrant population. This in indeed the case when the distribution in 2005 is deduced from Figure 2.2.

Figure 2.4. Male emigration propensities by duration of residence and average propensities across duration of residence

a) Non-Western immigrants *b) Western immigrants*



Note: For demographic events the duration of residence refers to the status at the end of the year. Emigration by duration of residence of d years in year t is applied to the population holding a duration of residence of d-1 years at the beginning of year t. Propensities indexed by a duration of residence of "5" are applied to the immigrant population with a duration of stay of 4 years and above at the beginning of the year. The depicted propensities are calculated as a three-year average of the actual propensities from 2009 to 2011. Individuals past the age of 70 are assumed to emigrate according the average propensity across duration of residence.

Source: Statistics Denmark.

That the propensities to emigrate from the population group comprising non-Western immigrants have been increasing for all categories of duration of residence might well be explained by the shift in the pattern of residence permits, cf. Figure 2.1, and the supposed relationship linking residence permits granted on the grounds of work or education to a higher emigration propensity than residence permits associated with asylum and family reunification. If work related residence permits are associated with a longer duration of stay than study related permits, this could offer an explanation as to why the propensity to emigrate has been decreasing for Western immigrants both within and across duration of residence.

Based on the above it is suggested that residence permit status is likely to be an important source to explain changes in demographic behaviour within a population group, and hence changes in the distribution of the immigrant population across duration of residence. Further, the historical development suggests that including duration of residence when describing emigration might still leave part of the behaviour unexplained, since the emigration propensity by gender, age, origin, and duration of residence is not constant over time. Behavioural changes therefore might still be attributed to differences in the grounds on which residence permits are warranted, leaving duration of residence an imperfect instrument for residence permit status. Data simultaneously characterizing an individual by duration of residence and by residence permit status are required in order to eliminate variation in demographic behaviour caused by the latter. Thus, though not favouring a forecast approach omitting emigration by duration of residence, the existence of unexplained variation might suggest that improving forecast accuracy perhaps requires more than simply expanding the present framework by an additional covariate.

3. The projection model

In cooperation with Statistics Denmark, DREAM is responsible of performing the official annual national projection of the Danish population. The projection is based on the DREAM population model, which is a group-based model determining the annual changes in the resident population stock from separately forecasted propensities of demographic events. Unlike an individual-based framework primarily known from microsimulation models, a group-based model projects simultaneously the behaviour of a group of individuals with identical demographic characteristics. Averaging numerous executions of a microsimulation model will result in a projection almost similar to the one resulting from a group-based model, given identical behavioural assumptions. The group-based model holds the advantage of executing faster than the individual based model, but suffers from memory drainage when adding further demographic characteristics to the model.

The latter feature constitute a significant restriction on introducing numerous additional population characteristics into the model framework. However, regardless of the model framework applied, introducing additional characteristics will reduce the number of observations with identical properties, hereby challenging the estimation of future demographic flows. Advanced data mining techniques can be applied in order to classify observations, leading to identical demographic behaviour being estimated across various characteristics. Introducing such an approach is reserved to further studies. In the projection model variation in demographic behaviour induced by lacking observations is predominantly reduced by using a simple three-year average of demographic flow propensities or by assuming identical demographic behaviour across origin groups.

Confidence bands indicating the uncertainty related to the estimation of demographic flows cannot be deduced since the projection model is entirely non-parametric.

3.1. The baseline model

The baseline model projects the development in the Danish population by gender, one-year age groups and origin given assumptions on future mortality, fertility, propensities to emigrate, immigrate, and change citizenship. The origin dimension consists of ten population groups defined by the combination of three characteristics: type of origin, country of origin, and citizenship status.

The type of origin categorizes an individual as a member of the immigrant, the descendant or the residual population, where the latter is by far the largest group. Immigrants and descendants are statistical concepts created by using specific rules in connection with information on family relations, citizenship and country of birth. According to Statistics Denmark, an immigrant is defined as a person born abroad whose parents are both foreign citizens or were both born abroad[6]. A descendant is a person born in Denmark whose parents are either immigrants or descendants with a foreign citizenship[7]. A person will be classified as a member of the residual population if just one of the parents is a Danish citizen and born in Denmark. The definition of origin type is exhaustive, meaning that all residents can be classified as a member of exactly one of these three groups. The country of origin identifies whether an individual stems from a Western or a non-Western country and is used to

[6] If there is no available information on either of the parents and the person was born abroad, the person is also defined as an immigrant.

[7] If there is no available information on either of the parents and the person in question is a foreign citizen, the person is also defined as a descendant.

subdivide immigrants and descendants[8]. Further classifying individuals according to Danish or foreign citizenship status results in the following ten population groups: The residual population with and without Danish citizenship, Western immigrants with and without Danish citizenship, non-Western immigrants with and without Danish citizenship, descendants of Western immigrants with and without Danish citizenship, and descendants of non-Western immigrants with and without Danish citizenship. Note that the information on citizenship status is necessary in order to categorize an individual as a member of the descendant population.

The projected population by age *x+1*, gender *g*, origin/population group *o*, at the beginning of year *t+1* ($N_{t+1}^{x+1,g,o}$) is determined from the population stock at the beginning of the previous year t ($N_t^{x,g,o}$) adjusted for immigration ($\overline{I}_t^{x+1,g,o}$ and $I_t^{x+1,g,o}$), emigration ($U_t^{x+1,g,o}$), change in citizenship ($\overline{S}_t^{x+1,g,o}$ and $S_t^{x+1,g,o}$), and number of deaths ($D_t^{x+1,g,o}$) during year *t*:

$$N_{t+1}^{x+1,g,o} = N_t^{x,g,o} + \overline{I}_t^{x+1,g,o} + I_t^{x+1,g,o} - U_t^{x+1,g,o} + \overline{S}_t^{x+1,g,o} - S_t^{x+1,g,o} - D_t^{x+1,g,o} \quad (1.1)$$

The number of persons aged 0, i.e. $N_{t+1}^{0,g,o}$, is determined from the number of births ($F_t^{g,o}$) and the previous mentioned demographic flows:

$$N_{t+1}^{0,g,o} = F_t^{g,o} + \overline{I}_t^{0,g,o} + I_t^{0,g,o} - U_t^{0,g,o} + \overline{S}_t^{0,g,o} - S_t^{0,g,o} - D_t^{0,g,o} \quad (1.2)$$

Note that the age dimension of the same cohort is dated differently depending on whether the population stock or demographic flows are considered. The age of the population is dated according to the age at the beginning of the year whereas the age of demographic flows refers to the age at the end of the year in which the demographic event occurred.

An overlined variable in equations (1.1) and (1.2) indicates, that the flow is determined exogenously in the projection model. Immigration to the two population groups consisting of immigrants from Western and non-Western countries without Danish citizenship, $\overline{I}_t^{x,g,o}$, is determined exogenously, leaving $\overline{I}_t^{x,g,o}$ empty for eight out of ten origin categories. The event of changing citizenship occurs within each of the following five groups: Western and non-Western immigrants, descendants of Western and non-Western immigrants, and the residual population. The number of individuals leaving a foreign citizenship, $S_t^{x,g,o}$, is determined endogenously. This simultaneously defines an equivalent exogenous increase in the number of individuals holding a Danish citizenship, hence $\overline{S}_t^{x,g,DANISH} = S_t^{x,g,FOREIGN}$. The model ignores the rare event of individuals acquiring a foreign citizenship[9].

---

[8] The group of Western countries consists of: all members of the EU, Andorra, Iceland, Liechtenstein, Monaco, Norway, San Marino, Switzerland, The Vatican, Canada, USA, Australia, and New Zealand. The non-Western countries are the remaining countries.

[9] This event can be related to the concept of dual citizenship where a person is simultaneously holding a Danish and a foreign citizenship until renouncing the former.

Most demographic flows are determined endogenously from a combination of an exogenous probability of an event occurring and the population exposed to risk, where the latter is unknown until the beginning of the projection year in question. Hence, the demographic flows in (1.1)[10] are determined according to

$$\begin{aligned} I_t^{x+1,g,o} &= i_t^{x+1,g,o} E_t^{x,g,o} \\ U_t^{x+1,g,o} &= u_t^{x+1,g,o} E_t^{x,g,o} \\ D_t^{x+1,g,o} &= d_t^{x+1,g} E_t^{x,g,o} \\ S_t^{x+1,g,o} &= s_t^{x+1,g,o} E_t^{x,g,o} \end{aligned} \tag{1.3}$$

where $E_t^{x,g,o}$ defines exposure to risk and $i_t^{x+1,g,o}$, $u_t^{x+1,g,o}$, $d_t^{x+1,g}$, and $s_t^{x+1,g,o}$ are the propensities to respectively immigrate, emigrate, die, and change citizenship during year *t* for all combinations of gender, age, and origin. Exposure to risk is defined as

$$E_t^{x,g,o} = N_t^{x,g,o} + \frac{1}{2}\bar{S}_t^{x+1,g,o} \tag{1.4}$$

where the latter part of the expression is only defined for individuals of Danish origin. Mortality and the propensities to immigrate, emigrate, and nationalize are all exogenous. By rearranging the terms of (1.3) the propensities can be calculated from historical data. The propensities to immigrate, emigrate, and nationalize are assumed constant in the projection and are calculated as a three-year-average of the propensities latest available. The Lee-Carter method is used to forecast mortality from data ranging from 1990 until the latest year available. Due to sparse observations on descendant mortality, the Lee-Carter estimation is only performed by gender and not by origin[11]. The number of births in (1.2) are determined from origin specific fertility rates depending on the age of the mother and a term of exposure to risk consisting of the female population in the middle of the year aged between 15 and 49. Fertility rates are exogenously determined from origin specific forecasts restricting total fertility by a predefined target of cohort fertility.

Consistent with the definition of an immigration event, cf. section 3.3, immigrants are not exposed to the risk of emigrating, dying or changing citizenship within the year of arrival. However if an immigrant on arrival is aged between 15 and 49, current fertility rates will apply.

Though the baseline model does not incorporate behaviour depending on duration of residence, the model is keeping track of the expected future composition of the immigrant population across duration of residence. By using the distribution of the immigrant population on duration of residence prior to the first projection year and by assigning a duration of residence of zero years to future

[10] The definition of exposure to risk for persons aged 0 is here ignored for convenience. DREAM (2013) can be consulted for clarification.

[11] The descendant population is typically aged below 40, requiring future demographic behaviour concerning older age groups to be imputed. Age-dependant fertility, propensities to emigrate, immigrate, and change citizenship are typically imputed from scaling the age profiles for the residual population. Imputation by scaling requires that a significant number of observations of the same demographic event is simultaneously available for descendants and the residual population. Considering that mortality is a rare propensity when aged below 40, this suggest that an approach ignoring the origin dimension is applied when estimating mortality.

immigrants on arrival[12], the duration distribution can be deducted at any time in the projection, which is a necessary condition when implementing duration dependant behaviour.

3.2. Introducing emigration by duration of residence

The projection model does not allow immigrant behaviour to depend on type of residence permits, though this is likely a significant source to explain variation in demographic behaviour within origin groups. Due to the data issues mentioned in section 2, duration of residence and not residence permit status is used to account for variations in demographic behaviour not explained by gender, age, and origin. However, as previously mentioned, variation within each duration group is still left unexplained. The model does therefore not project the future composition of residence permits. Due to data sparseness, the analysis is restricted to cover only re-emigration from the immigrant population from Western and non-Western countries holding a foreign citizenship.

The propensity to emigrate depending on duration of residence, *d*, is defined as

$$u_t^{x+1,d+1,g,o} = \frac{U_t^{x+1,d+1,g,o}}{E_t^{x,d,g,o}} \qquad (1.5)$$

where $U_t^{x+1,d+1,g,o}$ is the number of persons emigrating during year *t* from origin group *o*. As mentioned previously, *x* + 1, and *d* + 1 respectively refers to the age and duration of residence at the end of year *t*. Similarly to (1.4) exposure to risk, $E_t^{x,d,g,o}$, is defined as

$$E_t^{x,d,g,o} = N_t^{x,d,g,o} + \frac{1}{2}\bar{S}_t^{x+1,d+1,g,o} \qquad (1.6)$$

for *x*, *d* ≥ 0 [13]. Since, the analysis is restricted to cover re-emigration of immigrants holding a foreign citizenship, the last term of (1.6) can be ignored. As in the baseline model, the emigration propensities in (1.5) are calculated as a three-year-average of the propensities latest available. Benefitting from already tracking the duration distribution of the immigrant population in the baseline model, the expansion induced by introducing emigration by duration of residence, practically reduces to inferring and applying emigration propensities by an increased level of detail.

3.3. Data

Data on the resident population and the demographic events outlined above are available from Statistics Denmark´s Population Register, which is updated daily with information received from the Central Person Register. Data are preliminary since the daily updates also include information on

[12] Assigning a duration of residence of zero years applies only to immigrants holding a foreign citizenship. Immigrants holding a Danish citizenship are re-immigrating and are distributed on duration of residence according to the distribution of the resident immigrant population at the time of arrival. In this analysis, this distinction is of no relevance, since duration dependant behaviour will only be introduced for immigrants holding a foreign citizenship.

[13] For brevity, exposure to risk for persons aged 0 is not presented here. Individuals having just immigrated (the case of d = -1) are not exposed to risk of either emigration or change in citizenship. This restriction is imposed in order to ensure consistency with the rules identifying an immigration event in the data, cf. section 3.3.

previous events, which have not been timely reported to the Central Person Register[14]. Requiring information on citizenship restricts the time period covered to comprise the years from 1992 to 2012. Within this time period data categorizing the resident population by gender, single year of age and duration of residence, and origin are available. Data on demographic events can be obtained from 1992-2011.

The annual number of immigrants are defined as the number of individuals who are present at the end of the year, absent at the beginning of the year, and who have not been born in Denmark during the year. I.e. individuals immigrating and emigrating within the same year are not registered as immigration events in the data used for the population projection. The number of emigrants are defined as the number of individuals present at the beginning of the year, absent at the end of the year, and who have not died during the year. Individuals emigrating and then re-immigrating within the same year are not registered as emigrants[15].

Including duration of residence results in sparse observations on demographic events for certain combinations of demographic characteristics. This leads to the analysis being restricted to cover only re-emigration from the non-Western and Western immigrant population holding a foreign citizenship. Subjecting data on all demographic events to classification techniques imputing complete age profiles for all combinations of gender, age, and duration of residence is reserved to future studies. With the exception of emigration, immigrant behaviour is therefore subjected to the same assumptions as used in the baseline projection model.

Though available information on duration of residence ranges from 0 to 16+ years of stay, no significant variation in emigration behaviour can be identified above the fourth year of residency[16]. This suggests that no information will be lost by collapsing these categories, hence forming a new category of maximum duration of residence comprising data ranging from four years of stay and above.

According to (1.5) the propensity to emigrate for age a + 1 and duration d + 1 during year t is applied to the population exposed to risk characterized by age a and duration of residence d at the beginning of year t. E.g. the population exposed to risk with duration of residence = 0 at the beginning of year t is multiplied by the propensity to emigrate for duration of residence = 1 to establish emigration during year t from the group characterized by the shortest duration of stay. The immigrant population holding a duration of residence = 0 at the beginning of year t, has immigrated during year t - 1 and therefore has a duration of stay less than one year. Using five groups of duration of residence, emigration flows are categorized by 1-5+ years of residence meaning that the population exposed to emigration has resided in Denmark for 0-1, 1-2, 2-3, 3-4, and 4 years or more.

---

[14] Persons, who migrate, have to inform the local register in the municipality where they have residence. Failing to do so timely delays the registration of the migration event.

[15] Note that the definition of immigration and emigration used in the model is not equivalent to the definition used in the official migration statistics of Statistics Denmark. In the latter, migration is defined as an event implying a border crossing, i.e. a single individual will count as two immigration events if entering, exiting and re-entering within the same year. In the data used in DREAMs projection model this pattern would only represent a single immigration event.

[16] Though not statistically verified, introducing additional categories of duration of residence will not alter the results of section 4.

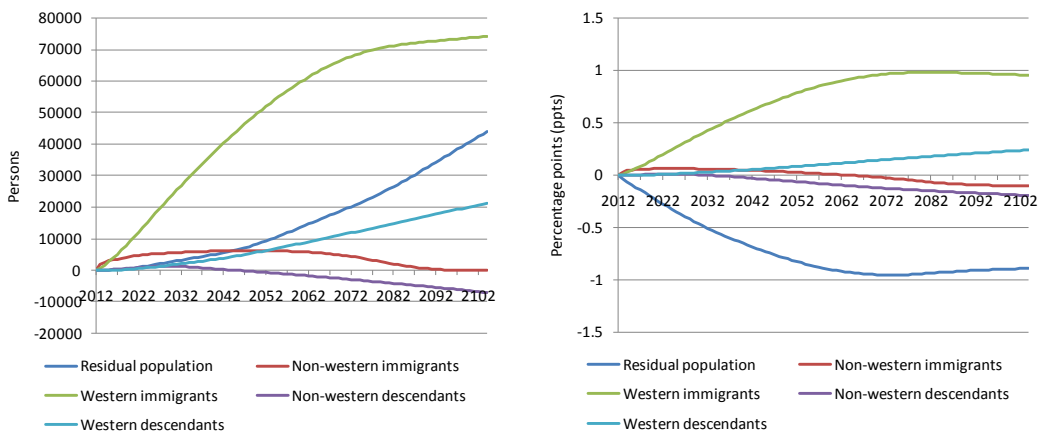
4. Results

Duration dependant emigration behaviour is introduced in the 2012-version of the DREAM population projection model[17]. As mentioned in section 3.1 and 3.2 the propensity to emigrate is assumed constant throughout the projection period and is calculated as an average of the propensities from 2009 to 2011. The expanded model projects the development in the resident population throughout the 21st century using the population at the beginning of 2012 as a starting point. This section will outline the implications of introducing emigration by duration of residence by comparing the composition of the resident population by origin and the distribution of the immigrant population on age and duration of residence to the baseline scenario.

Compared to the baseline projection introducing emigration by duration of residence leads to a gradual increase in the number of resident Western and non-Western immigrants. Throughout the entire projection period the immigrant population from Western countries is continuously larger than in the baseline scenario, whereas when reaching the end of the forecast horizon this is no longer the case for immigrants from non-Western countries. The change in the number of descendants from Western and non-Western countries will to a large extent mirror the development in the immigrant population. Further the resident residual population is increasing compared to the baseline scenario, cf. Figure 4.1.a. Western immigrants and descendants constitute a larger share of the total population compared to a projection omitting emigration by duration of residence, cf. Figure 4.1.b.

Figure 4.1. Change in the resident population by origin compared to the baseline scenario

a) Absolute change by origin *b) Change in distribution by origin*



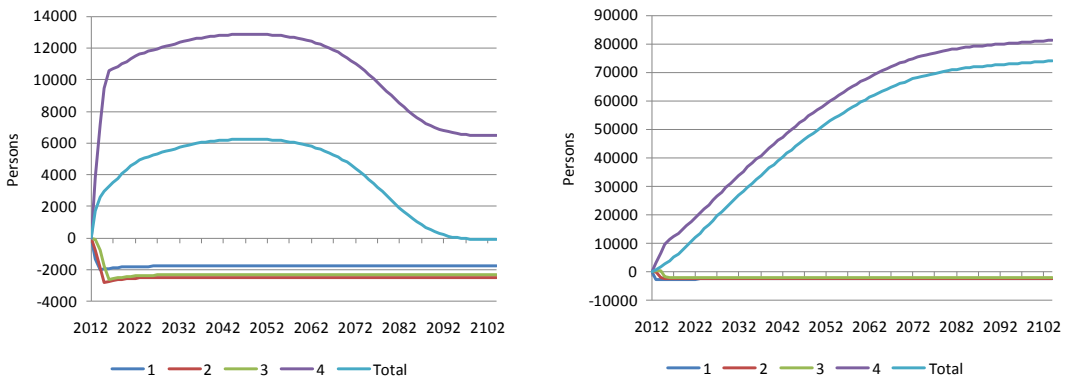
Source: The DREAM population projection model.

Observing the change in the immigrant population by duration of residence reveals that introducing duration dependant emigration has induced a decrease in the number of immigrants with a duration of stay less than or equal to three years, and an increase in the population with a duration of stay of four years and above, cf. Figure 4.2.

[17] Note that an analysis conducted on the 2013-version of the population projection, i.e. an analysis using updated assumptions on the underlying data and model framework, will lead to conclusions consistent with what is presented in this section.

Figure 4.2. Change in resident population by origin and by duration of residence

a) Non-Western immigrants *b) Western immigrants*



Note: Individuals having just immigrated are not exposed to the risk of emigration, hence the population holding a duration of residence of zero years is not subject to change.

Source: The DREAM population projection model.

This development is explained by the relationship between two effects: the direct effect and the dynamic effect. The direct effect is a composition effect which is defined as the effect obtained by combining the emigration propensities by duration of residence with exposure to risk by duration of residence. As depicted in Figure 2.4, the propensity to emigrate is typically larger than the average emigration propensity for individuals having immigrated within the most recent years, whereas the emigration probability lies below the average for the longest duration of stay. Since the share of immigrants in the initial population holding a long duration of residence is by far the largest, cf. Figure 2.2, introducing emigration by duration of residence will lead to an overall decrease in emigration, hence an increase in the resident immigrant population. This is subject to the assumption of a constant annual level of future gross immigration. The effect of a higher propensity to emigrate from the immigrant groups with short residencies is therefore dominated by the fact that the largest share of the immigrant population is facing a decrease in the propensity to emigrate.

Whether composition effects will lead to an increase or decrease in total emigration might vary between age groups depending on the distribution on duration of residence within the age group. If re-emigration from population groups with short residencies increases more in absolute terms than emigration from long residencies, the total population is immediately reduced. Implicitly the number of resident immigrants subject to future emigration in old age groups are reduced, hence reducing the direct effect of lower than average emigration propensities for long residencies in the future. Given the decrease in emigration from long residencies dominates within an age group, the total population increase will potentially enhance the composition effect within old age groups in the future. Effects evident across time are dynamic effects.

The change in both the Western and non-Western immigrant population results from the interaction between the direct and the dynamic effect. The interaction changes over time hereby explaining the differences in the nature of change in long residencies between Western and non-Western immigrants. In order to aid comprehension of the latter, age profiles of the changes in Figure 4.2.a and Figure 4.2.b are extracted for the years 2013, 2050, and 2100 respectively. From these profiles it

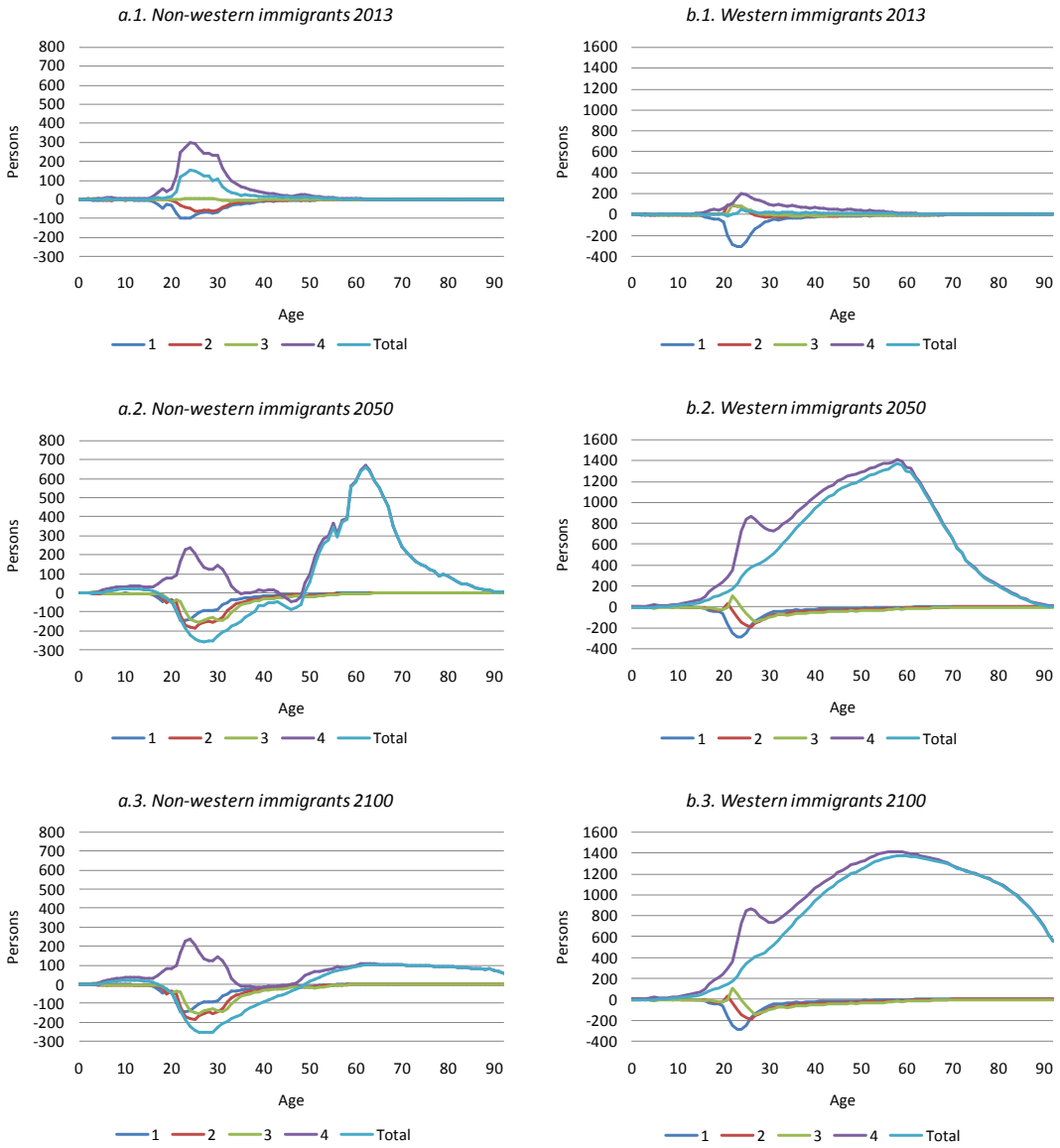
becomes evident that the nature of the change in the immigrant population compared to the baseline scenario can be explained by a shift in the distribution of the change on age and duration of residence over time.

The immediate direct effect of introducing emigration by duration of residence is depicted for the first forecast year 2013 in panel 1.a and 1.b of Figure 4.3. For both non-Western and Western immigrants, introducing emigration by duration of residence induces a decrease in the population holding the shortest residency, and an increase in the population holding the longest residency, overall resulting in an increase in the total population. Note that the composition effect at this point in time is far more evident for non-Western immigrants than for Western immigrants and that no dynamic effect is present.

Moving to 2050, the change in the immigrant population by age compared to the baseline scenario is qualitatively identical for Western and non-Western immigrants when comparing residencies lasting three years or less, cf. Figure 4.3, panel a.2 and b.2. Due to emigration propensities above average, the immigrant population with short residencies has decreased compared to the baseline model omitting duration dependant behaviour. Though the emigration propensities for short residencies are higher than the average emigration probability for most age groups, the age composition of the population change of short residencies is limited to ages below 30. This is explained by the fact that the number of immigration events, hence the number of resident immigrants holding short residencies, are rare above this age. For immigrants aged above 70, it is assumed that average emigration propensities apply, cf. Figure 2.4. Due to the fact that the share of immigrants holding short residencies later in life is rare, testing reveals, that the results of this section are robust to imposing average emigration behaviour to immigrants aged 45 or above.

When observing the change in the population groups of maximum residency there is now a remarkable difference in the age composition between Western and non-Western immigrants. For Western immigrants the population has increased significantly for all age groups resulting in a total increase in the population of Western immigrants. The population of non-Western immigrants holding the maximum residency has increased for ages 20-35 and for ages above 50 leading to an decrease in the immigrant population below 50 and an increase in the population above this age. However, as illustrated in Figure 4.2.a the result at this point in the projection period is an overall increase in the resident population of non-Western immigrants. The difference in the age composition of the change in long residencies between Western and non-Western immigrants is partly explained by differences in the variability between the average propensity to emigrate and the propensity to emigrate from the population group holding the longest residency. From Figure 2.4 it is evident that variability between the aforementioned propensities is larger for Western than non-Western immigrants suggesting that the direct effect of imposing emigration by duration is larger for Western immigrants, if assuming that the distribution on duration of residence is identical within both the Western and the non-Western immigrant population. The increase in the immigrant population aged above 50 is primarily explained by dynamics, i.e. the accumulation of direct effects similar to the ones evident in 2013. Contrary to the pattern of 2013, introducing emigration by duration of residence now induces a decrease in the total non-Western immigrant population aged below 50. Due to dynamic effects this will lead to a different future pattern of change in the population aged above 50 than observed in 2050.

Figure 4.3. The change in the immigrant population by age, origin, and duration of residence compared to the baseline scenario, selected forecast years



Source: The DREAM population projection model.

In 2100 the age profiles of the change in the resident population of Western and non-Western immigrants aged below 50 are practically identical to the pattern in 2050 for all residencies. However, a remarkable difference has occurred in the age groups above 50. The increase in the non-Western immigrant population has been reduced compared to 2050. Since the average propensity to emigrate for this age group almost coincides with the emigration propensity of the group holding the longest duration of stay, the impact of the direct effect is rather weak in the age groups above 50, cf. Figure 4.3 panel a.1, hence suggesting that the reduced change is almost entirely attributable to the dynamic effect. I.e. the effect of reducing the population with short residencies, and thus decreasing the population exposed to risk of emigration later in life, will dominate the immediate and direct effect obtained by applying lower than average emigration probabilities to individuals aged above 50. This

type of dynamics is exactly the consequence of the aforementioned shift from 2013 to 2050 in the age pattern below the age of 50. From 2050 to 2100 the age pattern of the change in the number of Western immigrants above the age of 50 has changed as well. Contrary to the non-Western immigrants, the change relative to the baseline scenario has increased during the projection period implying a continuing accumulation of population increasing direct effects in the ages below 50.

The difference in the age pattern describing the total change in the Western and non-Western immigrant population, aids explaining the nature of the change in the descendant population. Unlike the change in the number of Western descendants which mirrors the increase in the population of Western immigrants, the number of non-Western descendants decrease almost immediately compared to the baseline scenario despite the number of non-Western immigrants being larger, cf. Figure 4.1.a. As outlined in Figure 4.3, following an initial increase the total number of non-Western immigrants have generally decreased in the age group below the age of 50. This age group comprises the childbearing ages ranging from 15-49, hereby promoting a decrease in the number of non-Western descendants compared to the baseline scenario.

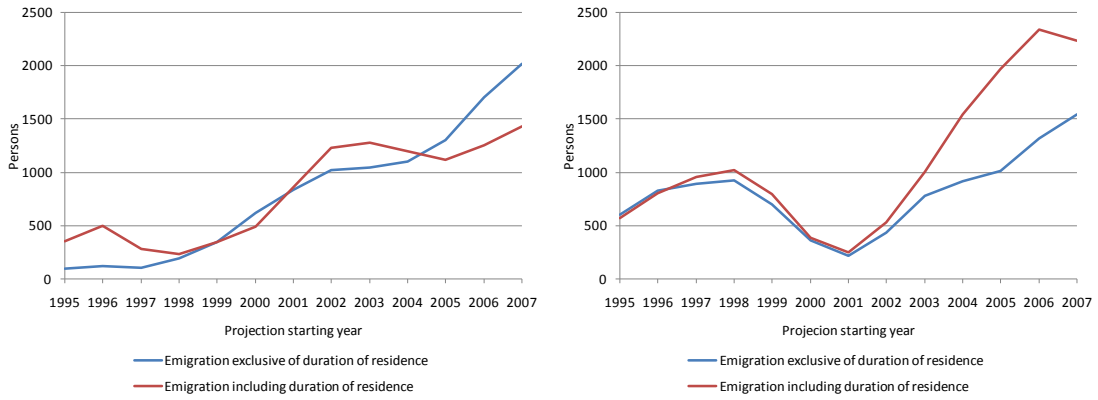
5. Assessing forecast precision

By performing a sequence of within-data population projections with a 5-year forecast horizon, it is established whether introducing emigration by duration of residence enhances forecast accuracy. Following the approach from above, emigration propensities respectively including and omitting duration of residence are calculated from a three-year average of the propensities prior to the starting point of each projection. Other demographic events are estimated by using the actual rates or propensities. Note especially that gross immigration to the population groups consisting of non-Western and Western immigrants with a foreign citizenship is exogenous and hence reflects the increase in annuals inflows during the recent decade. Data on emigration ranges from 1992-2011, hence the starting point of the first projection is the resident population by gender, age, origin, and duration of residence in the year of 1995. By estimating emigration propensities as an average of 1992-1994, the number of emigrants are projected for the years of 1995-1999 and the absolute difference between actual and forecasted emigration is used to assess forecast precision. Based on the population in 1996 and an average of the emigration propensities from 1993-1995, emigration is subsequently forecasted until the year 2000. By repeating this procedure annually through the starting year 2007, a total of 13 within-data projections are obtained for each assumption on emigration behaviour.

The result of the multiple forecasts is ambiguous since not uniquely identifying whether or not duration of residence should be considered when modelling emigration behaviour. For each of the 5-year projections respectively including and omitting emigration by duration of residence, Figure 5.1 depicts by projection starting year the average across forecast years of the absolute difference between actual and forecasted emigration.

Figure 5.1. Average annual difference between actual and forecasted emigration by projection starting year

a) Non-Western immigrants *b) Western immigrants*



Source: The DREAM population projection model.

For non-Western immigrants the dominant approach varies by projection starting year. E.g. forecasting emigration in the years 2002 to 2006 using the 2002 population as the projection starting point, will result in an average annual difference to actual emigration of approximately 1.250 persons when considering duration of residence and approximately 1.000 persons when using the baseline assumptions, suggesting that the latter approach should be preferred. Using 2006 as the projection starting year will lead to the opposite conclusion. For Western immigrants including emigration by duration of residence will almost consistently lead to the largest deviance from actual emigration. From 2003 and onwards the difference obtained by considering emigration by duration of residence becomes significantly larger than if using the baseline assumptions. E.g. using the 2005 population as a starting point and assuming duration dependant emigration will result in an average annual difference to actual emigration being almost twice the difference obtained if omitting duration of residence.

Though not apparent from Figure 5.1, forecasted emigration from the population of non-Western immigrants typically undershoot the actual emigration, whereas forecasted re-emigration from Western immigrants is typically overshooting the actual levels. This pattern of deviance is qualitatively identical within the projections respectively including and omitting emigration by duration of residence. In general this indicates that using a three-year average of the emigration propensities prior to the projection starting year induces poor forecast performance when changes in emigration behaviour occur within the forecast horizon. Since the emigration propensities have been increasing for non-Western immigrants, cf. Figure 2.3, undershooting the actual immigration is a comprehensible consequence of estimating emigration behaviour by a three-year average of the propensities prior to the projection starting year. The decrease in Western emigration propensities during the recent decade explains that overshooting is prevalent for this origin group.

The direct effect, i.e. the composition effect, aids explaining why using emigration propensities by duration of residence, will lead to a larger forecast inaccuracy than when using identical emigration behaviour across residencies. The increase in annual immigration contribute to an increasing share of immigrants holding short residencies. As mentioned in section 2, using constant emigration propensities by duration of residence is then likely to result in an increase in emigration because of

composition effects regarding duration of residence. Such effects are absent when using the baseline assumptions, thus reducing the effect of over- or undershooting.

6. Conclusion

During the recent decade changes in the composition of the resident immigrant population across duration of residence is largely attributable to changes in annual immigration quantities and changes in immigrant behaviour. Previous studies linking resident permit status with duration of residence, suggest that the former is an important source when explaining changes in immigrant behaviour, and hence changes in the distribution of the immigrant population on short and long term residencies.

Recognizing that emigration behaviour varies across duration of residence, motivates that introducing this characteristic when forecasting emigration will eliminate some of the variation in emigration patterns not already accredited to gender, age, and origin.

Initially the consequences of introducing emigration by duration of residence in the baseline model was investigated. Under the assumption of constant exogenous annual immigration and constant re-emigration propensities, composition effects inducing a decrease in the average emigration propensity lead to an increase in the immigrant population. Using duration dependant emigration propensities further causes a change in the age distribution of immigrants, subsequently resulting in a change in the size of the descendant population.

A series of within-data projections respectively omitting and including duration dependant emigration were conducted in order to establish whether emigration by duration of residence, will improve forecast accuracy. However, partly due to the presence of changing emigration behaviour within each category of duration of residence, the historical development in emigration patterns can not entirely be dedicated to composition effects obtained by combining the resident population by duration of residence with constant emigration behaviour by duration of residence. Applying emigration by duration of residence is then more likely to over- or undershoot actual emigration than baseline emigration behaviour. This suggests that emigration by duration of residence should be applied with caution if expecting behavioural changes within population groups characterized by gender, age, origin and duration of residence. I.e. though immediately appealing, increasing the number of covariates describing demographic behaviour combined with the assumption of constant future behavioural propensities, constitute a significant risk of over- or undershooting actual behaviour due to composition effects.

Alternatively the subject of section 5 could have been approached by statistically validating the performance of the projection models respectively including and omitting emigration by duration of residence. Having selected the best model, prediction error could subsequently have been assessed using a different data set. This approach involving the use of training and test data does not only encourage an assessment of whether or not emigration behaviour should depend on duration of residence, but also invites to address if a constant three-year average is to be preferred when estimating emigration propensities. Further it can be verified if the use of additional covariates

describing the decision to emigrate enhances model performance. Subjecting the baseline assumptions to various competing approaches is reserved to further studies[18].

Due to data sparseness, the analysis restricted the introduction of duration dependant behaviour to emigration. Data mining techniques can aid the imputation of incomplete age profiles by classifying observations across characteristics, allowing duration of residence to be applied to e.g. fertility rates and the propensity describing the change in citizenship[19]. This extension is also an obvious subject for future studies.

7. References

Hansen, M. F. & Stephensen, P. (2012): Danmarks fremtidige befolkning - Befolkningsfremskrivning 2011. Rapport, marts 2012.

Statistics Denmark (Danmarks Statistik) (2008): Indvandrere i Danmark 2008, December 2008.

Statistics Denmark (Danmarks Statistik) (2011): Indvandrere i Danmark 2011, December 2011.

[18] Competing approaches used to estimate emigrate propensities could potentially comprise classification models, logit and probit models.

[19] Introducing duration dependent changes in citizenship would potentially improve the definition of the population group exposed to risk of emigration. However, since most changes in citizenship status happens after four years of residency, notable effects would demand more than five categories of residency.