University of Copenhagen
Institute of Economics
6 Studiestraede, Dk-1455 Copenhagen K
Denmark

Phone *+45 35 32 32 61*
Fax *+45 35 32 32 59*

E-mail ***HansOluf.Hansen@econ.ku.dk***
Home page http://www.econ.ku.dk/personal/usihoh/
Ftp-server ftp://ftp.ibt.ku.dk/usihoh/

2001-01-24

Modeling External Migration, Reproduction, Citizenship and Ethnic differentials In Labor Supply[1]

By

Hans Oluf Hansen & Karsten Längerich

Abstract

The issue of population heterogeneity related to differentials in external migration, fertility and labor supply is currently subject to considerable public and political concern in Denmark and elsewhere. The knowledge base of the Danish debate at this point largely consists of partial ad-hoc studies, mostly founded on official statistics and extracts in terms of more or less casual joins of existing population registers. Notwithstanding questionable data quality, access to unpublished official population data tends be overly bureaucratic and exceedingly costly. So how do we add more rigors to interpretation of complex social issues imperfectly mirrored in available population data? To reconcile the dilemma of clarifying a complex population matter with the statistically insufficient data at hand we address a model proposed by Hansen and Gustafsson (1997). We consider a state space featuring known central heterogeneity aspects regarding external migration, fertility and labor market activity and an associated parameter space of intensities in continuous time. A model application in terms of three scenarios enables us to reach consistent, rich, and very interesting social conclusions from the specified assumptions.

[1] Published in Proceedings of Symposium on Applied Statistics, arranged by Univ. of Copenhagen, Dept. of Econ. and Statistics Denmark , January 22-24, 2001.

1 Introduction and background

The welfare consequences of below-replacement reproduction are subjects to growing concern in many economically developed countries. In most Central and Western European countries this development is concurrent to influx and long-term establishment of new ethnic groupings, sometimes of rather untraditional national origin. Denmark has experienced below-replacement reproduction on a cross-sectional basis over the past three decades. Moreover, since the early 1960's the country has passed from an ethnically and socially fairly homogeneous nation to an emerging multi-ethnic society (Matthiessen 2000; Hansen 2000a). To some circles this development exhibits a threat to national identity and cherished traditional cultural values.

Concerning the welfare consequences the crux of the matter could perhaps be summarized as follows: How does below-replacement reproduction and external migration affect labor supply and employment, and thereby the so-called 'dependency burden' faced by the gainfully employed fraction of the resident population? Can the adverse socio-economic effects of the on-going population aging be mitigated by deliberate policies promoting, for example, integration and favoring immigration of persons with attractive qualities in terms of human or physical capital? What are the ceteris-paribus impacts for the growth and structure of the Danish population of observable ethnic differentials in reproduction, mortality, labor force mobility, and external migration at the turn of the millennium? And which demographic impacts will eventually emerge from an indefinite continuation of a given multiethnic regimen of fertility and mortality and a constant number of immigrants with unchanging composition by sex, age, and foreign ancestry?

To arrive at coherent conclusions and thoughts based on informed understanding of the growth and structure of a complex population process of great importance and interest to many people, adequate conceptualization and appropriate modeling is strongly required. The knowledge base of the Danish debate at this point to a large extent consists of partial ad-hoc studies, mostly founded on official statistics and extracts in form of more or less arbitrary joins of existing population registers. Production of official statistics by second-hand use of administrative data files obviously gives rise to a wide range of problems in demographic application. This is true, for example, of non-reported censoring and ad-hoc defined concepts such as *descendants* and artificial constructs like *degree of unemployment*. Important processes in continuous time, for instance labor force participation and employment are commonly recorded in discrete time and reported on an aggregated basis. Aggregated gross flows in terms of linked discrete-time observations of segments of event histories are only published on rare occasions.

So how do we add more rigors to analysis of complex social issues imperfectly mirrored in available population data? To clarify the analytical potentials and limitations of accessible data and existing concepts we need an appropriate theoretical framework. Some basic analytical concepts and categories are identified in Section 2. We address the modeling issues in section 3. The project is summarized in section 4.

2 What are the relevant analytical concepts?

According to Randall Kennedy**,** a professor at Harvard Law School, *integration* is best understood in relation to two other key words, namely *segregation* and *desegregation*[1]. Ethnic and racial desegregation can safely be taken for granted to be the principal goal and objective of the Danish constitution (*Grundloven*), of the Government, and of the Parliament (*Folketinget*). On the other hand, it should not be forgotten that the fundamental rights of lawful Danish residents include freedom to assemble, freedom of conscience and worship, and full proprietary rights. The Constitution rules out forced desegregation. Social integration of smaller ethnic groups within a larger national population therefore has many dimensions. First of all, what is *ethnicity*? And how do we make ethnicity operational in demographic analysis and consistent prediction of socially relevant multidimensional population dynamics? The Question Bank, a resource From The Centre for Applied Social Surveys (CASS) in England, offers the following definition:

> '*An ethnic group is a collectivity within a larger population having real or putative common ancestry, memories of a shared past, and a cultural focus upon one or more symbolic elements which define the group's identity, such as kinship, religion, language, shared territory, nationality or physical appearance. Members of an ethnic group are conscious of belonging to the group'* [2]

Extensive social research published in refereed scientific journals over the years has indicated *time* in terms of *sojourn as resident of a country* to be an influential variable associated with the level of ethnical desegregation. The theme of the present study is particularly concerned with ethnicity related to *nationality* or *citizenship* and to native and foreign *ancestry*. Identifying current legal citizenship of any lawful member of the resident population would rarely present a problem in developed countries. Categorization by *nationality* or *citizenship* therefore has a long and established standing in official population statistics worldwide. This was true also of official Danish population statistics prior to 1991. However, according to Professor P.C. Matthiessen, a scientific

[1] Source: http://www.igc.apc.org/dissent/archive/summer96/kennedy.html
[2] Source: http://qb.soc.surrey.ac.uk/nav/fr_home.htm

advisor of Statistics Denmark for several decades, official statistical categorization of the immigrant population by citizenship has become too narrow for integration policies and other purposes: (1) 'Many naturalized former immigrants will not have been integrated in the Danish society'; (2) 'The children of naturalized immigrants will not belong to the sub-population of immigrants'; and (3) 'The propensity to become Danish citizens by naturalization will differ across the nationalities' (Matthiessen 2000, pp. 92-93). To remedy such demands, a new classification has been introduced in Danish official population statistics, namely that of *immigrants, descendants* (Da. *'efterkommer'*) and *others*. Referring to Statistics Denmark (1998) Matthiessen (2000, p. 93) notes that '*immigrants* are Danish residents born abroad of parents that are both foreign nationals or born abroad' while '*descendants* are Danish residents born in Denmark of parents either being immigrants or descendants'[3]. Detailed reviewing of the absurd consequences of this descendant definition is not possible within the framework of this paper, nor shall we consider all other - not always logically disjoint - formal definitions of the concepts 'descendants' and 'others' used in Danish official population statistics over the past decade.

To not a small degree at the expense of the traditional classification by citizenship, the new creative sooner than clarifying categorization is being used for two types of purposes, namely production of official statistics (based, primarily, on the national population register CPR) and official population projection. The resulting statistical tables production demonstrates *de facto* the data-driven character of the varying definitions.

The conceptual and methodological innovation in Danish official population statistics entails description and analysis of a time-continuous and duration-dependent population process by cross-sectional classification based on ad-hoc defined fixed and time-dependent personal covariates. Let alone the logically and definitional shaky efforts of making *ethnicity* operational along with the obvious shortcomings of a cross-sectional scheme of observing a complex time-continuous population process with duration dependence: What is the quality of the basic data source(s) on which the statistics is based? The basic source - the CPR Register - was initiated on April 1, 1968. Its restricted c*overage in time* and a significant number of *defective or missing embedded pointers* between parental couples and their offspring set up tight limits for computerized genealogical research to identify members of the ethnic groupings applied by Statistics Denmark. The previous event histories and genealogies of the resident population born before April 1, 1968, are frequently incompletely mirrored in the CPR,

[3] As far as the definition of *descendants* is concerned, Statistics Denmark (1998, p. 11) is misquoted or misunderstood by Matthiessen (2000, p. 93).

probably more so as regards resident foreign nationals than Danish nationals with ancestry born in Denmark over many generations.

3 Modeling external migration, reproduction, citizenship and ethnic differentials in labor supply

On elucidating ethnical integration dynamics the expediency of the approach adopted by Statistics Denmark around 1991 remains to be seen for the reasons given in Section 2. The behavioral changes indicated by the published empirical statistics refer to incompletely observed event histories in terms of aggregated outcomes of sample paths in complex finite state spaces of time-continuous two-sex human branching processes anchored in time. How are each of the processes associated with the multi-dimensional population process at large? What are the state space(s)? What are the associated parameter spaces? And which types of applications have been produced up to now? Due to the official character of the definitions and the modeling approaches of Statistics Denmark, we start by reviewing briefly an ethnic population projection published by Statistics Denmark (1997). We then review a multi-dimensional model proposed by Hansen & Gustafsson (1997) referring to an extensive application of the latter model under Markov conditions, see Längerich (1999)[4].

3.1 Modeling external migration, reproduction, and ethnic differentials (Statistics Denmark 1997:16)

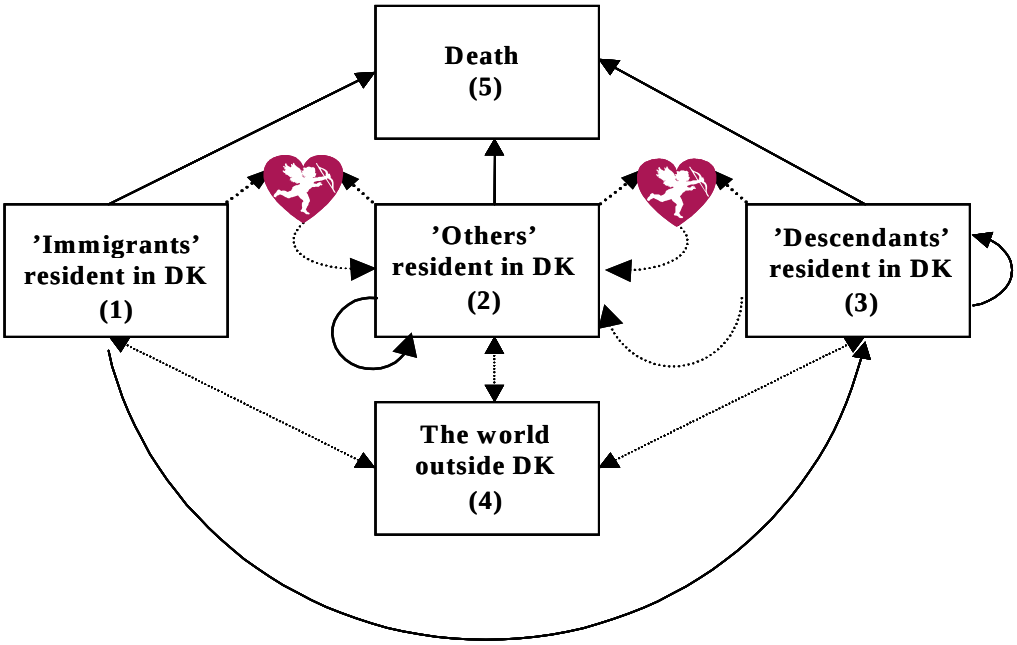
An official population projection reported by Statistics Denmark (1997:16) describes 'possible courses of development up to the year 2020' regarding the number of immigrants and descendants with origin from economically developed nations or economically developing countries of the world.

State space An explicit outline of the state space does not appear to be available. The flow diagram exhibited in Statistics Denmark (1997:16, p. 7) connotes to computational algorithms and flows of controls, e.g. represented by a computer program, rather than to transitions in a state space of a time-continuous stochastic process.

From the flow diagram we identify the state space shown in figure 1. The three resident sub-populations communicate with the world outside Denmark via net migration. The only communication between the sub-populations is entry in terms of live birth. The arched curves denote exogenous/endogenous entry by live birth into the sub-

[4] **Download:** ***ftp://ftp.ibt.ku.dk/usihoh/docs/Students/Laengerich.pdf***

Figure 1 Modeling population development by ethnic ancestry[1]



1) Based on Statistics Denmark 1997:16, figure 1, p. 7.

groups 'Others' and 'Descendants'. A part of the exogenous entry by live birth to 'Others' is generated by un-modeled two-sex biological liaisons indicated by the uni-directional dotted lines leading to the hearts. The notions of *nationality* and *naturalization* are also latent in the model. The hidden two-sex mechanism governs the distribution of the immigrant live births on 'Others' and 'Descendants' and - together with latent naturalization of female descendants - the allocation of the descendants' offspring on 'Others' and 'Descendants'.

The two-sex mechanism implied by the definitions of ethnic groupings used in the 1997-projection by Statistics Denmark is brought to operate via exogenous ad-hoc multiplier(s) based on historical evidence as of January 1, 1995 through 1997. The decision rules are as follows: A live birth is added to 'Others' (1) If a woman belonging to the group of 'Immigrants' or 'Descendants' begets a child by a male belonging to the group 'Others'; (2) If a *naturalized* female descendant begets a child - no matter the ethnicity of its father (Statistics Denmark 1997:16, p. 5).

Introducing a two-sex mechanism without proper modeling does appear a dubious and very dangerous strategy with regard to measuring past and current demographic experience concerning population heterogeneity associated with ethnicity and as a rational basis for coherent anticipation of future population development. One ironic consequence of the un-modeled two-sex mechanism being that future population development will almost surely make historically based fixed multipliers outdated very

fast. The realism of the multipliers is crucial for the future population distribution on the categories 'immigrants', 'descendants' and 'others'.

Parameter space The parameter space associated with the state space in figure 1 eventually boils down to three separate sets of intensities or hazard functions in terms of central population rates. Such intensities may be multivariate in character, although not in the application of Statistics Denmark up to now. Net migration and the above-mentioned ad-hoc multipliers are exogenous to the model and therefore to the associated parameter space(s).

Applications The latest model application concerns projection of the number of resident immigrants and descendants up to the year 2020 based on a dichotomous division of the world outside Denmark. The application should be seen 'as computational examples showing the consequences of the chosen assumptions' (Statistics Denmark 1997:16, p. 15). According to Matthiessen (2000, p. 93) this projection ' makes it possible to obtain an impression of the development over the first twenty years of the 21st century.' The authors of the present paper would rather warn against such creative and methodologically unclear computations dressed up as an official statistical population projection. Such calculations appeal more to individual feelings and subjective interpretation than to reason and common logic.

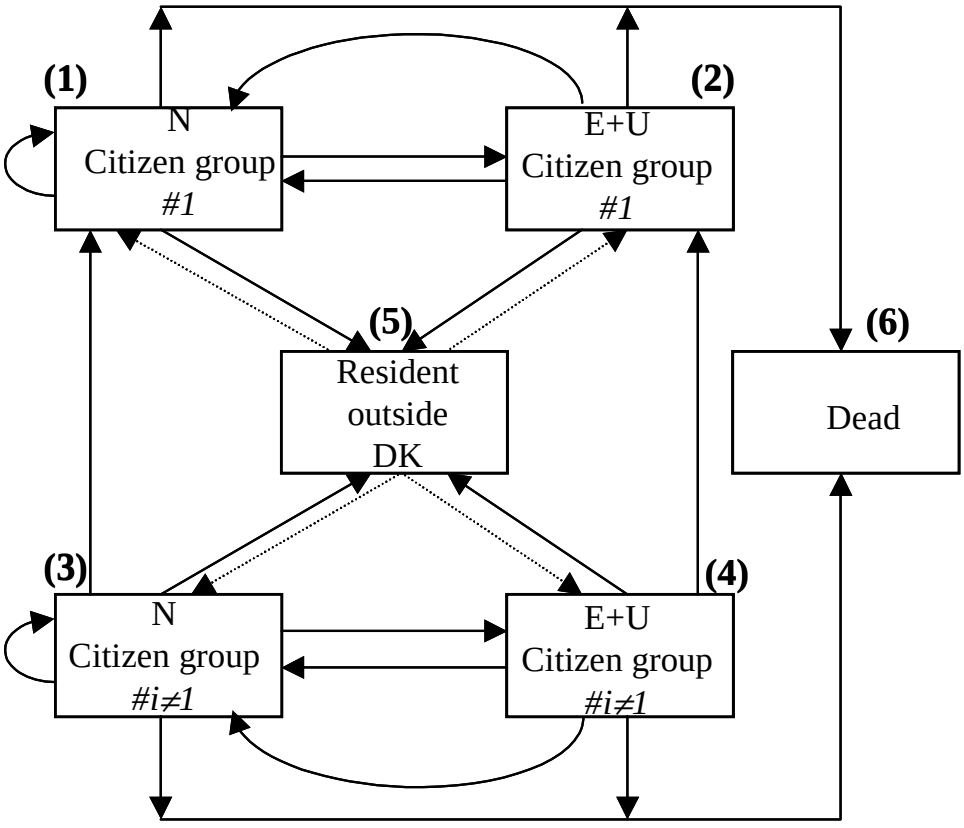
3.2 A model of external migration, naturalization, reproduction and labor force participation (Hansen & Gustafsson 1997)

How does below-replacement reproduction and external migration affect labor supply, and thereby the so-called 'dependency burden'? To provide consistent answers to such a question we need appropriate demographic multi-state modeling of relevance not only to this problem, but also to existing knowledge and available population data. Sections 2 and 3.1 have shown that ethnicity related to demographic differentials is intrinsically connected with nationality and citizenship, and possibly with changing living conditions and behavioral adaptation that may take generations to unfold. Fortunately we do not have to choose between categorization by *citizenship* and by *ethnical origin and progeny* on modeling population heterogeneity associated with external migration, reproduction, mortality, citizenship, ethnic differentials, and labor supply. Hansen & Maxim (1996) and Hansen (1996) propose a simple demographic multi-state model of external migration, citizenship, naturalization, reproduction and mortality that provides a much better description of such population heterogeneity than the two-state life models applied, for example, by Statistics Denmark (see Section 3.1), Feichtinger & Steinmann (1992) and Steinmann (1994). In the multi-state model

members of any non-Danish citizen group are eligible to become Danish nationals by naturalization while resident in Denmark. External migration, reproduction, and mortality may be differential by citizen group. The model allows of distinguishing between naturalized and non-naturalized offspring of a given resident citizen group, either on a cross-sectional or a generational basis. Hansen & Gustafsson (1997) extended the model to accommodate labor force mobility in continuous time and - to some extent - *ethnical origin and progeny* for each sub-group of citizens under study.

State space The state space of the extended model by Hansen & Gustafsson (1997) is shown in figure 2. Like with figure 1 the arched curves refer to endogenous or exogenous entries in terms of live birth while the dotted straight lines refer to (exogenous) immigration. Immigration and emigration are separate events that may depend on nationality and labor force status. A newborn live child acquires the citizenship of its mother. Danish nationality may be obtained by naturalization at any age with an intensity that may be contingent e.g. on personal labor force status and nationality. The events of naturalization and change of labor force status are taken to be non-coinciding over the individual life courses.

Figure 2 A model of differential labor force participation, differential external migration, and differential reproduction among citizen groups resident in Denmark



Legend,
N = Not in the workforce; E=Employed; U=Unemployed

Parameter space The parameter space associated with the state space in figure 1 consists of a set of time-continuous intensities (hazards) of any complexity, for example multivariate and time-dependent in character. Under Markov assumptions the associated parameter space $\underline{\underline{M}}(x)$ of the state space exhibited in figure 2 looks as follows. Let *x* denote *age* and let $m^{ij}(x)$ be a piece-wise constant function of age i.e. $m^{ij}(x) = m^{ij}(x+\tau), \tau \in [0,t[$ with *t* indicating length of the age interval starting in *x* and the top scripts indicating transition types in the state space. Furthermore, let fertility in life state *i* be given by $f^{i}(x) = f^{i}(x+\tau), \tau \in [0,t[$. The intensity matrix $\underline{\underline{M}}(x)$ is then,

$$\underline{\underline{M}}(x) = \begin{pmatrix} -\sum_{j=2}^{6} m_x^{1j} & m_x^{12} & 0 & 0 & m_x^{15} & m_x^{16} \\ m_x^{21} & -\sum_{\substack{j=1 \\ j\neq 2}}^{6} m_x^{2j} & 0 & 0 & m_x^{25} & m_x^{16} \\ m_x^{31} & 0 & -\sum_{\substack{j=1 \\ j\neq 3}}^{6} m_x^{3j} & m_x^{34} & m_x^{35} & m_x^{36} \\ 0 & m_x^{42} & m_x^{43} & -\sum_{\substack{j=1 \\ j\neq 4}}^{6} m_x^{4j} & m_x^{45} & m_x^{46} \\ 0 & 0 & 0 & 0 & -\sum_{\substack{j=1 \\ j\neq 5}}^{6} m_x^{5j} & 0 \\ 0 & 0 & 0 & 0 & 0 & -\sum_{j=1}^{5} m_x^{6j} \end{pmatrix}$$

The probability matrix $\underline{\underline{P}}(x)$ may be established by straightforward exponential series expansion of the intensity matrix i.e.

$$\underline{\underline{P}}(x,t) = \sum_{n=0}^{\infty} \frac{t^n}{n!} \cdot \left(\underline{\underline{M}}(x)\right)^n$$

Using the **P**-matrix, the row vector ***I(x,t)*** (referring to age-state specific exogenous immigrant flows), and the initial multistate population represented by the row vector **B,** projection in the model may then be accomplished by the following (slightly pragmatic) operation,

$$\underline{B}^{T_1}(x+t,t) \cong \underline{B}^{T_0}(x,t) \cdot \underline{\underline{P}}(x+\frac{t}{2},t) + \underline{I}^{T_0}\left(x+\frac{t}{2},t\right)$$

The available population data, mostly in terms of published official population statistics, do not allow advanced hazard estimation. They are, however, informative enough

to be statistically sufficient for partial estimation of some, but not all of the forces of transition in Markovian applications of the model.

The population process modeled in figure 2 contains many instances of obvious duration dependence. With more sophisticated hazard specification, including duration dependence, the population process would no more be Markovian. Mathematically impeccable estimation and prediction in the model based on a Semi-Markovian version of the state space may be accomplished by Monte Carlo simulation in the manner of Hansen (2000b, 2000c).

Applications So far there exists only one application of the model in figure 2, namely that of Längerich (1999). This report is available at the web site listed in footnote 4. The application gives an idea about the ceteris-paribus consequences of external migration, reproduction, mortality, citizenship and ethnic differentials in labor supply in Denmark on the turn of the millennium. The set of assumptions underlying the application has been subject to severe restrictions imposed by inadequate data availability. Starting off from empirical prevalence rates in terms of labor force participation rates and using a three-state life model applied by Hoem (1977) under Markov assumptions, 'guestimates' of the latent intensities of labor force accession and withdrawal from the labor force were obtained as elements belonging to a set of intensity solutions complying with the observed prevalence rates.

4 Summary and conclusion

Can immigration alleviate the adverse economic consequences of reproduction below replacement? If the volume of external immigration steadily is large enough, the total size of the resident population may well be upheld, even at a higher level than we know of today. However, if the immigrated residents naturalize and rather quickly adopt a sustained below-replacement reproduction like that of the 'native resident population' - whatever that is - a certain amount of population ageing in the ensuing stationary resident population would then become persistent as well. The economic feasibility of a future population development of this nature would very much depend on the qualities of human and physical capital of the immigrants and of the costs of integrating the immigrants into gainful economic activities in the Danish society.

References

Feichtinger, Gustav and Günther Steinmann (1992). *Immigration into a population below replacement level. The case of Germany*. **Population Studies 46**, 275-284.

Hansen, Hans O. (2000a). *A Review of Some Current Danish Population Trends.* Chapter 5 in Kucera, T. and E. Schaich (2000, eds.), **New demographic face of Europe** (Springer Verlag)

Hansen, Hans O. (2000b). *An AIDS Model With Reproduction - With An Application Based on Population Data from Uganda*. **Mathematical Population Studies 2000, Vol. 8(2)**,pp. 175-203.

Hansen, Hans O. (2000c). *Lecturing Notes - Demography 2* (Unpublished)

Hansen, Hans O. (1996). *Some Demographic Impacts of Current Regional and Ethnic Differentials in Denmark.* In M.C. Debernardi (1996, ed.), **Evolution or Revolution in European Population, Vol. 2**, 129-144 (Milano: Franco Angeli).

Hansen, Hans O. and Lars Gustafsson (1997). *Analysis of the Danish Labor force mobility 1985 - 1996.* Paper prepared for the ESPE 97 Annual Conference, ESRC Research Centre, University of Essex, England, 18-21 June, 1997 (Unpublished)

Hansen, Hans O. and Paul S. Maxim (1996). *Some Patterns and Social Impacts of External Migration on a Below-Replacement Population: Denmark by Turn of the Millennium.* **Scandinavian Population Studies 11**, 136-149 (Helsinki: The Nordic Demographic Society).

Hoem, Jan M. (1977). *A Markov Chain Model of Working Life Tables*. **Scandinavian Actuarial Journal** 1977, 1-20.

Längerich, Karsten (1999). *En modelbaseret fortolkning af udviklingen i den danske befolkning i første halvdel af det 21. århundrede med fokus på etniske grupper og disses deltagelse i arbejdsstyrken (A Model-Based Projection of Danish Population Development in the First Half of the 21st Century Focused on External Migration, Citizenship and Ethnic differentials in Labor Supply*, 44 pages (Univ. of Cph., Dept. of Econ.: 6-points essay in DEMOGRAPHY 2) (Unpublished).

Matthiessen, P.C. (2000). *Indvandringen til Danmark I det 20. århundrede (Immigration to Denmark in the 20th century).* **Nationaløkonomisk Tidsskrift 139 (2000),** pp. 79-94.

Statistics Denmark (1997*).* ***Befolkning og valg nr. 1997:16*** *(Population and General Elections*) (Description of model and assumptions)

Statistics Denmark (1998). ***Indvandrere i Danmark*** *(Immigration to Denmark)* (Cph.: Danmarks Statistik)

Steinmann, Günther (1994, ed.). ***Population Consequences of Population Change in Industrialized Countries*** (Berlin: Springer Verlag)