European Journal of Population **18:** 387–409, 2002.

Modelling Regional Immigration: Using Stocks to Predict Flows

NICOLE VAN DER GAAG* and LEO VAN WISSEN
*Netherlands Interdisciplinary Demographic Institute (NIDI), P.O. Box 11650, 2502 AR The Hague, the Netherlands (*Author for correspondence: e-mail: gaag@nidi.nl)*

Received 26 November 2001; accepted in final form 22 May 2002

Van der Gaag, N. and Van Wissen, L., 2002, Modelling Regional Immigration: Using Stocks to Predict Flows, European Journal of Population, **18**: 387–409.

Abstract. This paper explores the feasibility of improving regional international migration assumptions by analysing the relationship between international migration flows and foreign population structures. Regional projection models are discussed and empirically tested using data for Sweden. The results show that regional assumptions on international migration of foreigners could be improved by using the spatial distribution of stocks of foreigners as predictor. Improvements, however, are only minor and do not seem to compensate for the loss in simplicity of the models. This is especially true at the aggregate level of immigrant groups. Only for relatively new immigrant groups, improvements are substantial.

Key words: international migration, regional projection models, stocks of foreigners, Sweden

Van der Gaag, N. et Van Wissen, L., 2002. Migration internationale par région: comment prendre en compte les stocks pour estimer les flux, Revue Européenne de Démographie, **18**: 387–409.

Résumé. Cet article cherche à améliorer les estimations de migrations internationales par régions en analysant la relation entre les flux migratoires et les structures de population étrangère. Les modèles régionaux de projection y sont discutés et testés sur les données suédoises. Il apparaît que les estimations régionales de migrations internationales pourraient être améliorées par la prise en compte de la distribution spatiale des étrangers. Cependant ces améliorations restent modestes et ne compensent pas la complexité des modèles qu'elles requièrent. Ceci est particulièrement vrai au niveau global de grands groupes d'immigrants. Les gains en précision ne deviennent substantiels que pour les nouvelles vagues d'immigrants.

Mots clés: migration internationale, population étrangère, projection régionale, modèles de projection, Suède

1. Introduction

In recent history, international migration has become increasingly important as a source of population change. In a growing number of countries in the European

Union (EU), international migration is now more important than natural increase (Münz, 1996; Van der Gaag et al., 1999). Despite its key role in population growth, however, migration is very difficult to project. The uncertainties surrounding migration are tremendous as migration is often related to historical events and depends heavily on national policies.

In the past, inflows of migrants have led to the settlement of large non-native groups in the European Union. Along with this development, the role of networks in international migration processes has gained importance (Hugo, 1981; Boyd, 1989; Fawcett, 1989; Wilpert, 1992; Massey et al., 1993; Böcker, 1994; Esveldt et al., 1995). In general, the existence of migrant networks in the destination country may reduce moving costs for potential migrants of the same origin. Waldorf (1996) provides a detailed theory of the possible effects of the migrant population on potential immigrants. She distinguishes, following Gurak and Caces (1992), selective and adaptive forces. The *selective* forces attract migrants from the origin. Selectivity is the result of the established social network, and has the effect of providing information to potential migrants, lowering moving costs for newcomers and facilitating new moves through family reunification. With every migrant, the social network increases, thereby propelling new immigration, until the pool of potential migrants is drained and the system becomes saturated. Moreover, as time passes, the average duration time of the migrant stock increases. These compositional changes have profound effects on the selective capabilities of the stock. A common hypothesis is that concentration occurs with new immigrants, followed by later dispersion. *Adaptive* functions of the network are important for the likelihood of return migration. The larger the network, the larger its adaptive function (short term assistance, longer term integration, etc.) and the lower the return propensity of newcomers. Return flows, in turn, affect the composition of the migrant stock, and hence its selective functions.

Since networks of migrants may play a role in attracting new migrants, the question arises as to whether one can use information on stocks of foreigners to predict immigration flows. At the national level, immigration is associated with unemployment, but the links with social networks are less obvious from the empirical point of view (Van der Gaag and Van Wissen, 1999). Migrant groups, however, tend to be highly concentrated in just a few – often urban – regions. Most likely, therefore, the links between migrant networks and the size of immigration flows are more evident at the regional level.

This paper explores the feasibility of improving the quality of international migration assumptions at the subnational level by analysing the relationship between international migration flows and foreign population structures. The ultimate aim is to develop and improve methods to analyse and project the regional allocation of international migration flows using stocks of foreign residents as predictor.

2. Regional projection models of international immigration

In this section, regional projection models of international immigration are discussed from a methodological point of view. A commonly used method to produce population projections is the cohort-component method. Following this approach, the complete population structure, i.e., the way in which the total population is classified by gender, age group, marital status and so on (cohort), can be projected over a future period of time. In the simplest form, the base population of sex (s) and age group (x) is transformed into a projected population of sex (s) and age group ($x + 1$) by taking explicitly into account the components of population change. Typical components are mortality, fertility and international migration (Van Imhoff et al., 1994). Within the framework of the standard cohort-component projection model, international migration is a difficult component. In such a projection model, the components birth and death are modelled as events that occur in a population exposed to the risk of experiencing these events. An occurrence-exposure rate can be calculated by taking the number of events that occurred in the unit time interval and dividing it by the total exposure time experienced by the population. These rates are calculated for each age and sex-combination. Emigration can be treated similarly. However, *immigration* does not fit in this framework. The population at risk is very large and heterogeneous (i.e., the population living in the rest of the world) and therefore the risk of immigration into the country is very small. Calculating rates based on these occurrence-exposure intensities will in general not be feasible. Instead, other methods are usually employed.

Basically there are three methods for projecting international migration (Van Imhoff et al., 1994):

1. Models that produce estimates of net migration totals;
2. Models that produce separate estimates of emigration and immigration totals;
3. Models that produce separate estimates of emigration rates (by age and sex), and immigration totals.

Net migration, and emigration and immigration totals are usually broken down into age- and sex-specific categories using a sex-specific age profile. In practice it is possible that the projection model uses net migration figures, but that the underlying migration assumptions are based on separate immigration and emigration hypotheses and models.

In *regional* population projections, the same distinction applies, and usually the same choice of method is made at the national and the regional level. In addition, a basic distinction is made between regional models with and without consistency with national projections. Consistency may be achieved in various ways (Van Imhoff et al., 1994: 63):

1. Following the bottom-up approach, in which the national projection is the aggregate of the regional projections;
2. The top-down approach, in which the results of the national projections are distributed over the regions, using an allocation rule;

3. A mixed approach, in which the regional results are adjusted in such a way that they add up to the national total.

The method used may be different for each component in the projection model. In general, regional immigration is treated in a top-down framework, where the projected national total immigration (by age and sex) is allocated over the regions, using regional shares.

Methods of calculating immigration shares may be characterized by the type of information they require. We distinguish three different methods. Models using *historical shares* are based on observed immigration flows over the regions in previous years. Immigration shares using *regional stocks* are based on the observed distribution of the (sub)population under study over the regions. Finally, methods using *additional (demographic or non-demographic) variables* to calculate regional immigration shares are often based on a regression-type equation that relates observed characteristics of the regions to observed regional immigration flows. In practice, mixed or hybrid forms do exist.

In this paper we discuss four models: models using (1) historical shares; (2) historical shares and stocks of foreign population; (3) non-linear stocks of foreigners and (4) non-linear stocks and additional variables. The models discussed are nested models, i.e., the structure of the simpler model is captured by the structure of the more complex models. In other words: starting with the simplest model (historical shares), additional parameters will be added to constitute a more complicated model.

HISTORICAL SHARES

The "standard" or most common approach in allocating immigration to regions is using historical information on the destination choices of migrants. The simplest form is to use the most recent regional shares, Q_r or a combination of shares of recent years, and apply these to the projection:

$$Q_r(t) = Im_r(t-1)/Im_+(t-1) \tag{1}$$

where $Q_r(t)$ is the regional share in immigration of region r at time t, $Im_r(t-1)$ is the size of the immigration flow into r at time $t-1$, and $Im_+(t-1)$ is the total immigration at the national level at $t-1$. Often, not only information from time $t-1$ is used but from a range of years $t-u$, for $u = 1, \ldots, U \leq t$ (for instance, a weighted average of previous years). An example of the historical share approach is given by Edmonston and Passel (1992), who apply it to different ethnic immigrant groups in the United States, but in practice in many European countries this is the common procedure (Van Imhoff et al., 1994), although in general immigrants are not broken down into different groups by ethnicity/nationality/origin. This method has also been applied in the most recent regional population scenarios of Eurostat (De Jong and Visser, 1997). If age- and sex-specific information is available, the shares can be age- and sex-specific as well.

The method of historical shares, however, suffers from a number of drawbacks. First, there is no direct theoretical foundation other than "inertia of the system" (today's migration pattern is very similar to yesterday's), and second, the historical shares are fixed whereas from a theoretical point of view the shares may change as a result of changes in the size and composition of the migrant stock through ageing, return migration, internal migration, and the influence of exogenous variables (political, economic). What makes the historical shares approach so popular is the limited amount of information that it requires. Information from one or a few years of regional immigration is sufficient.

HISTORICAL SHARES AND STOCKS OF FOREIGNERS

Migrant populations tend to be largely attracted to a limited number of urban areas and different migrant populations show different regional patterns (White, 1993; Coleman, 1994; Bucher, 1996; Van Huis and Nicolaas, 1999; Van der Gaag and Van Wissen, 2001). Most likely, therefore, the projection of the regional allocation of immigrants may be improved by including information on stocks of foreigners in allocation models.

One step towards a method incorporating stocks is a mixed approach of stocks and historical flows. This method is essentially expressed in the following equation:

$$Q_r(t) = (S_r(t)/S_+(t)) * F_r(t-1) \tag{2a}$$

where $Q_r(t)$ is, as before, the regional share of immigrants in region r at time t, $S_r(t)$ is the size of the stock in region r at the beginning of the projection period, $S_+(t)$ is the size of the stock at the national level at time t, and $F_r(t-1)$ is a regional immigration factor that specifies the deviation of the immigration shares, based on observed flows at time $t-1$, from the proportional allocation according to the shares of the stock. Thus,

$$F_r(t-1) = [Im_r(t-1)/Im_+(t-1)]/[S_r(t-1)/S_+(t-1)] \tag{2b}$$

When applied to total immigration, the stock variable S pertains to the total population or related quantity. For instance, in the Netherlands in the early 1990s a method was applied where S pertains to the housing stock, and F is a regional factor that specifies how the recent immigration pattern deviates from the proportional allocation according to the housing stock (Leering and Den Otter, 1992). Alternatively, the total population could have been used for S. The necessary input for this approach differs only from the historical shares method by S_r (the total population) by region, which is always available. As different migrant populations show different regional patterns, the allocation of immigration may be further improved by applying the model to each migrant group separately. If the relationship between stocks and flows is perfectly proportional, then $F = 1$ for all regions. In general, however, the factor is different from 1 (Van der Gaag et al., 2001).

MODELS USING NON-LINEAR STOCKS OF FOREIGNERS

Although model (2) takes into account the effect of stocks, it does not allow for variations in the size of the stock over time. F_r is fixed and therefore the non-proportionality of each region is constant over time. If, due to saturation effects such as discussed by Waldorf cited earlier (Waldorf, 1996), the law of diminishing marginal increase of attracting and retaining power applies, F should be made dependent on the size of the stock. This method has been applied in Huisman and Van Wissen (1998) where the following specification was used in a regional projection model for the Netherlands:

$$Q_r(t) = S_r(t)^{\alpha} / \Sigma_j S_j(t)^{\alpha} \qquad (3)$$

where Σ_j is the summation over all regions, $j = 1, \ldots, R$, such that $\Sigma_j Q_j = 1$. The power function S^{α} makes the non-proportionality of each region dependent on the size of the stock. Since S is endogenous in the projection model $Q_r(t)$ is endogenous as well, and captures the effects of the size of the migrant network, as well as some of the non-linear (i.e., saturation) effects and the fixed effect of historical shares. If $\alpha < 1$, then larger regions receive a less than proportional share of immigrants, if $\alpha > 1$ they receive a more than proportional share. If $\alpha = 1$ immigration is exactly proportional to the regional distribution of the stocks. In general, however, α should be smaller than 1, since saturation effects are likely to exist at the regional level.

The coefficient α may be estimated using generalized linear models (GLM's) and specifying the model in log-linear form. By doing so, however, estimated standard errors are too small, since the underlying assumption of the Poisson distribution, and hence equality of mean and variance, does not apply. Consequently, in this specification a correction for overdispersion is needed (for an application of this correction method, see Van Wissen and Visser, 1998). Huisman and Van Wissen (1998) applied this model in regional projections of the population according to origin. Their study showed that care should be taken in applying the method to regions with a very small stock. If α is smaller than 1, very small regions receive relatively large shares of immigrants. For one region having a size of only a fraction of the size of the largest urban regions, the relative impact of the projected large share of immigration turned out to be unrealistically high.

MODELS INCLUDING ADDITIONAL VARIABLES

Of all demographic components, migration interacts most directly with other domains of society, such as the labour or the housing market, as well as with political and institutional factors. Therefore, additional, often non-demographic, external characteristics may be useful information in predicting the regional distribution of international migration. On the other hand, external variables can usefully be employed in projection models only if these variables themselves can be predicted with a fair level of accuracy. For instance, unemployment may be an important indicator of regional shares, but the problem of predicting the future level of unemployment is as large as that of foreign immigration. These types of variables may be useful in population scenarios, where alternative economic developments are assumed and their consequences for population growth are studied. Urbanization variables, and other highly predictable characteristics of the regions may in principle be useful indicators. However, the net effect of static variables on regional shares is also captured by specifying the non-linear relationship between stocks and flows. Consequently, in projection models the usefulness of external factors is limited and restricted to time-varying but predictable variables. Nevertheless, in regional immigration scenarios these external variables are highly necessary because it increases the transparancy of the assumptions used (which is in line with Rees et al.'s (2001) recommendation about improving migration projections by incorporating non-demographic variables).

The functional form of a model using regional stocks and additional external variables is an extension of model (3). In a loglinear form this model is written as follows:

$$\log Im_r(t) = \alpha \log S_r(t) + \mathbf{X}_r(t)\boldsymbol{\beta} \tag{4a}$$

where $\mathbf{X}_r(t)$ is a vector of external variables for region r with values pertaining to the beginning of the projection period (1 January of year t), and $\boldsymbol{\beta}$ is a vector of coefficients to be estimated, as well as the coefficients α. Equation (4a) turns out to be equivalent to the multinomial logit model for grouped data:

$$Q_r(t) = \frac{S_r^{\alpha}(t)\exp[\mathbf{X}_r(t)\boldsymbol{\beta}]}{\sum_{j=1}^{R} S_j^{\alpha}(t)\exp[\mathbf{X}_j(t)\boldsymbol{\beta}]} \tag{4b}$$

A particular application of model (4) is the use of internal migration variables as predictors for regional immigration. As internal migrants are often directed towards different destinations from those attracting immigrants, this raises the question as to how immigration is related to internal migration. Frey and Liaw (1998) and Stillwell et al. (1999) address two competing hypotheses: the substitution hypothesis where vacancies on the housing and labour markets, as a result of outmigration of domestic residents to other regions, are filled by immigrants, versus the "immigrant push" hypothesis, where internal outmigration of domestic residents is caused by high immigration. If initial concentration of migrant groups will be followed by later dispersion, internal outmigration might be reinforced due to secondary migration of the immigrants themselves. Taking internal migration as predictor of regional immigration is in line with the first hypothesis and is based on the assumption that negative net migration leaves dwelling space open which makes immigration more attractive, while positive net migration has the opposite effect.

3. Regional models of immigration empirically tested

We have tested the models presented in the previous section using data for Sweden for the period 1988–1999 (source: migration and population statistics; Statistics Sweden). This test is restricted to immigration of foreigners; immigration of Swedish nationals are not taken into account. What is required is a breakdown of the international immigration flows and resident population according to region r, and immigrant group g. The exact definition of these immigrant groups is a matter of choice. Generally there is no agreement on the most appropriate statistical definition of this group. Different concepts are used, such as citizenship, country of birth, origin or some other definition of ethnicity. In this study immigrant groups have been classified according to citizenship (stocks as well as flows of migrants). A complicating factor of using citizenship is that we have to take into account naturalisations. The number of naturalisations in Sweden is rather high, varying from 3.5 to 9 per cent of the total stocks of foreigners in the given period, with a total number of 37.7 thousand in 1999. At the subnational level, however, differences in naturalisation propensities are only minor and correlations between regional shares in stocks of foreigners and naturalisations are rather high. Therefore we may assume that, although a definition of stocks of foreigners based on citizenship may underestimate the size of the migrant population, it can nevertheless be used as a good indicator of the regional distribution of foreigners. Since the data on regional stocks are from the population register, other changes in the size of the regional stock, due to internal migration, birth and death are automatically taken into account.

To test the regional models of immigration presented in the previous section, we estimated the parameters of the models 3 (non-linear stocks) and 4 (non-linear stocks and internal migration) using observed stocks and flows for 1992 as well as average stocks and flows over the period 1989–1992. Subsequently we used the results of the estimations to predict the regional pattern of immigration in the period 1993–1999. As the process of migration is not homogeneous across different groups of migrants, the models were tested for several (clusters of) nationalities, more or less following the classification expressed by Borgegård et al. (1996):

- All non-Swedish citizens
- EU-other (EU non-nationals); this group consists mainly of labour immigrants from other EU countries like Germany, France and the United Kingdom
- The aggregates of non-EU citizens
- Finnish citizens as a representative of unregulated neighbouring country immigration: Nordic citizens have almost unlimited freedom of movement between the Nordic countries and labour immigration in Sweden recruited from Finland has a long tradition. Although the share of Finnish immigrants in Sweden is declining, in 1999 still 10 per cent of all non-Swedish immigrants were Finnish citizens.

Table I. Immigration by citizenship, Sweden, 1988 and 1999

	1988		1999	
	Absolute	Percentage	Absolute	Percentage
Non-SE	44490	86.7	34572	69.4
EU-other	10624	20.7	8835	17.7
Non-EU	33866	66.0	25737	51.6
Finland	5422	10.6	3379	6.8
Iran	8142	15.9	1022	2.1
Iraq	1334	2.6	5526	11.1
South-America	3875	7.6	1094	2.2
Swedish nationals	6809	13.3	15266	30.6
Total immigration	51299	100	49838	100

- Iranian and Iraqi citizens as representatives of regulated non-European refugee immigration. These two nationalities form relatively new groups of immigrants in Sweden. While the immigration share of Iranians reached its highest level in the second half of the 1980s, inflows of Iraqis increased gradually until 1999. 16 per cent of the non-Swedish inflow in 1999 was formed by Iraqi immigrants. From 1984 to 1994 these immigrants were dispersed over the entire country, in accordance with the "Whole of Sweden" strategy. This strategy is aimed at mitigating the financial and social burden for individual municipalities and to create a possibility for sparsely populated municipalities to attain a more balanced population structure. Long-term effects, however, are unclear, as in time, many immigrants may move to other, often larger centers (Borgegård et al., 1996).
- South American citizens as a representative of unregulated non-European refugee immigration. A large part of this group is formed by immigrants from Chile.

In Table I an overview is given of immigration by citizenship in 1988 and 1999, respectively.

The data have been collected at the NUTS 3 level. The NUTS classification (Nomenclature of Territorial Units for Statistics) is a regional division of the European Union at different geographical scales. Each country (the NUTS 0 level) is divided into one or more NUTS 1 regions, which are divided into one or more NUTS 2 regions, and so on. For Sweden, there are 21 NUTS 3 regions (the 1998 NUTS classification), corresponding to the national administrative division "Län".

MODELS ESTIMATED

Since the exact relationship between internal and external migration is not known, we estimated different variants of model (4) that vary according to the exogenous information used: (4.1) using net internal migration; (4.2) using internal outmigration; and (4.3) using the log of internal outmigration. In all models aggregate internal migration figures were used, not specified by citizenship. The models have been specified and estimated as log-linear models. The value of the log likelihood function can be used for testing nested models, while the adjusted R^2 measures the success of the estimated model in predicting the value of the dependent variable, taking into account the number of independent variables in the model (contrary to the R^2, the adjusted R^2 may decrease as one adds independent variables which do not contribute to the explanatory power of the model).

Table II presents the results of the model estimations for base year 1992, and Table III for base period 1989–1992. The main results may be summarized as follows:

Base year 1992

- Models including stocks clearly outperform the null-models (models with a constant term only): the log likelihood increases significantly for all models tested. In terms of predictive power (adjusted R^2), models for Iran and Iraq were slightly less satisfactory compared to the other groups.
- With the exception of the models for Finland and EU non-nationals (of which a large part is formed by Finnish immigrants), the coefficient of the stock is considerably smaller than one. Therefore, regions with a relatively sizeable population of foreign descent generally receive a smaller than proportional share of immigrants.
- Adding internal migration variables to the models does not improve the estimations for the aggregate of all migrant groups. For the individual groups (Finland, Iran and Iraq) a slight improvement can be observed, mainly by including *net* internal migration. However, while the relationship between net internal migration and international immigration is positive for Finland, it is negative for Iranian and Iraqi immigrants. This might be related to regulations. While refugees may be allocated to those regions where dwelling space is left open, labour migrants may be more attracted to those regions where work is available, i.e., the same regions *Swedish* residents are attracted to.

Base period 1989–1992

- Estimations based on the averages of the period 1989–1992 are highly similar to the results based on observations for 1992 only. The effect of adding internal migration to the estimation models is even smaller. In general, in terms of adjusted R^2, model estimations for the period 1989–1992 are somewhat less convincing than estimations based on 1992.

Table II. Estimation results, Sweden, NUTS 3, base year 1992 (stocks as per 31 Dec 1991)

	constant		log(stock)		4.1: net mig 4.2: outmig 4.3: log outmig		log likelihood	
	C	z	alpha	z	beta	z	L	adj. R^2
Non-SE								
Null-model	**7.54**	27.44					–19554.8	
Model 3	–0.36	–1.39	**0.80**	32.66			–436.5	0.993
Model 4.1	–0.33	–1.12	**0.79**	27.70	5.58E-06	0.22	–435.6	0.993
Model 4.2	0.18	0.27	**0.73**	9.47	1.14E-05	0.87	–422.4	0.993
Model 4.3	*–1.77*	–2.19	**0.59**	5.01	*3.95E-01*	1.83	–381.4	0.992
EU-other								
Null-model	**5.68**	14.18					–5360.4	
Model 3	**–3.46**	–7.25	**1.00**	21.02			–283.2	0.977
Model 4.1	**–2.99**	–7.50	**0.94**	23.06	**1.35E-04**	3.75	–195.2	0.989
Model 4.2	**–3.33**	–2.50	**0.99**	5.74	3.02E-06	0.11	–283.1	0.975
Model 4.3	*–3.26*	–2.25	**1.03**	5.77	–4.95E-02	–0.15	–283.0	0.976
Non-EU								
Null-model	**7.37**	28.98					–14895.7	
Model 3	0.24	0.89	**0.76**	28.44			–435.1	0.987
Model 4.1	0.10	0.34	**0.77**	25.11	–2.61E-05	–0.98	–417.3	0.988
Model 4.2	*1.01*	1.76	**0.66**	9.39	1.79E-05	1.50	–397.0	0.992
Model 4.3	*–1.77*	–2.20	**0.48**	4.45	**5.20E-01**	2.61	–339.3	0.991
Finland								
Null-model	**4.86**	10.50					–2788.6	
Model 3	**–3.88**	–6.54	**1.01**	16.50			–215.5	0.960
Model 4.1	**–3.38**	–6.23	**0.95**	16.45	**1.78E-04**	2.61	–172.6	0.974
Model 4.2	**–3.83**	–2.95	**1.01**	5.73	1.13E-06	0.04	–215.5	0.956
Model 4.3	*–2.76*	–2.15	**1.14**	7.97	–2.44E-01	–0.97	–208.1	0.961
Iran								
Null-model	**5.15**	25.12					–1372.4	
Model 3	–0.02	–0.05	**0.70**	13.18			–196.4	0.920
Model 4.1	–0.46	–1.17	**0.76**	15.04	**–1.23E-04**	–2.74	–159.0	0.950
Model 4.2	–0.27	–0.37	**0.74**	6.54	–7.86E-06	–0.43	–195.2	0.914
Model 4.3	0.27	0.19	**0.74**	4.48	–6.05E-02	–0.22	–196.1	0.915
Iraq								
Null-model	**5.20**	24.97					–1463.6	
Model 3	0.77	1.60	**0.74**	9.90			–283.0	0.799
Model 4.1	0.09	0.17	**0.86**	10.85	**–1.56E-04**	–2.71	–217.2	0.856
Model 4.2	**1.54**	2.50	**0.57**	4.76	*3.15E-05*	1.76	–252.8	0.886
Model 4.3	–1.24	–0.92	**0.53**	3.58	3.72E-01	1.59	–258.1	0.868

Table continues on next page.

Table II. Continued

	constant		log(stock)		4.1: net mig 4.2: outmig 4.3: log outmig		log likelihood	
	C	z	alpha	z	beta	z	L	adj. R^2
South-America								
Null-model	**4.36**	10.93					–1401.4	
Model 3	**–1.55**	–4.12	**0.85**	18.04			–119.4	0.984
Model 4.1	**–1.53**	–3.68	**0.85**	15.57	8.60E-06	0.15	–119.3	0.983
Model 4.2	0.03	0.05	**0.54**	4.27	**6.19E-05**	2.54	–102.8	0.988
Model 4.3	**–5.95**	–3.30	**0.51**	3.69	**7.56E-01**	2.51	–102.7	0.978

Null-model: including a constant only (no additional variables).
Significant coefficient values; italic: p > 95% one tailed; bold: p > 99% one tailed.

- Contrary to the estimation based on 1992, the relationship between international immigration and net internal migration is negative for Finland, although not significant.

To summarize, we may conclude that models including information on stocks of foreigners successfully describe regional patterns of immigration flows. The differences between the fit of model 3 and any of the models 4.1 to 4.3 is not sufficiently large (if better at all) to justify the inclusion of internal migration variables into the models. Models based on observations for base period 1989–1992 perform slightly worse than models based on observations for 1992 only.

MODELS VALIDATED

In the next step of the study we investigated how well the models perform in predicting the regional distribution of future immigration flows. We compared the predicted regional immigration pattern for Sweden for the period 1993–1999 with the observed pattern for the same groups of immigrants. The models have been estimated for immigrants with the following citizenships: Finnish, EU non-nationals, South American, Iranian, Iraqi, non-EU and non-Swedish. We tested models based on base year 1992 as well as base period 1989–1992. For each group the following models have been tested: (1) historical immigration shares; (2) historical shares and regional stocks of foreigners; (3) non-linear stocks; (4.1) non-linear stocks and net internal migration; (4.2) non-linear stocks and internal outmigration; and (4.3) non-linear stocks and the log of internal outmigration. In addition, for the period 1989–1992, moving averages have been calculated to estimate regional immigration shares.

Table III. Estimation results, Sweden, NUTS 3, base period 1989–92 (stocks as per 31 Dec 1988–91)

	constant		log(stock)		4.1: net mig 4.2: outmig 4.3: log outmig		log likelihood	
	C	z	alpha	z	beta	z	L	adj. R^2
Non-SE								
Null-model	**7.75**	29.94					–22101.5	
Model 3	0.37	1.35	**0.75**	28.70			–613.9	0.989
Model 4.1	0.36	1.25	**0.75**	27.19	2.65E-06	0.13	–613.4	0.989
Model 4.2	0.61	0.95	**0.72**	9.78	4.26E-05	0.41	–609.0	0.987
Model 4.3	–1.11	–1.23	**0.54**	4.28	*4.03E-01*	1.72	–540.4	0.987
EU-other								
Null-model	**6.04**	15.98					–7181.2	
Model 3	**–2.71**	–5.01	**0.96**	17.72			–462.3	0.961
Model 4.1	**–3.20**	–5.77	**1.01**	18.13	**8.26E-05**	2.34	–376.2	0.972
Model 4.2	*–2.84*	–1.99	**0.98**	5.42	–2.31E-06	–0.10	–462.1	0.959
Model 4.3	*–3.18*	–1.87	**0.90**	4.35	1.12E-01	0.29	–460.4	0.958
Non-EU								
Null-model	**7.56**	31.93					–15933.5	
Model 3	**1.07**	4.02	**0.70**	26.04			–543.2	0.984
Model 4.1	**1.12**	4.30	**0.70**	26.40	–2.74E-05	–1.37	–500.3	0.984
Model 4.2	**1.53**	2.95	**0.64**	10.10	9.36E-06	1.03	–518.7	0.984
Model 4.3	–0.73	–0.83	**0.47**	4.20	*4.46E-01*	2.12	–453.8	0.987
Finland								
Null-model	**5.14**	10.74					–3824.4	
Model 3	**–4.03**	–7.77	**1.05**	19.81			–209.9	0.969
Model 4.1	**–3.70**	–5.12	**1.01**	12.63	–5.42E-05	–0.65	–206.3	0.973
Model 4.2	*–2.52*	–2.05	**0.85**	5.12	2.79E-05	1.25	-207.9	0.972
Model 4.3	**–3.17**	–2.77	**1.15**	8.89	–1.90E-01	–0.83	-204.6	0.971
Iran								
Null-model	**5.41**	32.60					–1260.7	
Model 3	**1.25**	2.58	**0.58**	9.08			-289.9	0.821
Model 4.1	**1.31**	2.77	**0.57**	9.15	–5.89E-05	–1.18	–273.9	0.811
Model 4.2	0.90	1.17	**0.64**	5.29	–9.97E-06	–0.60	–285.6	0.813
Model 4.3	2.22	1.45	**0.69**	3.77	–2.04E-01	–0.66	–284.7	0.813
Iraq								
Null-model	**4.76**	31.63					–655.5	
Model 3	**1.74**	4.24	**0.54**	7.92			–203.3	0.765
Model 4.1	**1.79**	4.75	**0.54**	8.48	*–1.07E-04*	–2.08	–176.3	0.818
Model 4.2	**1.53**	2.70	**0.60**	5.16	–9.34E-06	–0.57	–201.0	0.735
Model 4.3	**3.62**	2.49	**0.72**	4.86	–3.26E-01	–1.34	–191.5	0.742

Table continues on next page.

Table III. Continued

	constant		log(stock)		4.1: net mig 4.2: outmig 4.3: log outmig		log likelihood	
	C	z	alpha	z	beta	z	L	adj. R^2
South-America								
Null-model	**5.06**	15.96					–2024.1	
Model 3	0.32	1.24	**0.71**	21.11			–147.6	0.980
Model 4.1	*0.47*	1.78	**0.69**	19.69	–4.89E-05	–1.48	–138.6	0.982
Model 4.2	**1.30**	2.98	**0.51**	6.24	**3.35E-05**	2.52	–126.9	0.986
Model 4.3	*–2.57*	–1.90	**0.50**	4.93	*4.85E-01*	2.19	–130.1	0.986

Null-model: including a constant only (no additional variables).
Significant coefficient values; italic: p > 95% one tailed; bold: p > 99% one tailed.

As our measure of closeness of fit between predictions and observed distributions, we used the Relative Absolute Error (RAE), which is calculated as:

$$RAE = \frac{\sum_i |Obs_i - Exp_i|}{\sum_i Obs_i} \tag{5}$$

where Obs_i is the observed number of immigrants in region i, and Exp_i is the expected number as predicted by the model. This statistic measures the relative size of the summed absolute difference between observed and expected values. It has a lower limit of 0 (perfect fit), and is in most applications much less than 1. If the RAE is close to 1 this indicates that the total absolute error is of the same order of magnitude as the total observed inflow, i.e., a bad fit. Larger values than 1 are also possible, when the total absolute error is much larger than the total observed inflow. This indicates an extremely bad prediction.

Table IV and V present the summary results of the closeness of fit between predictions and observations for all groups of immigrants for base year 1992 and base period 1989–1992, respectively. These results indicate the following:

- In general, best predictions are found for the total group of non-Swedish nationals. RAEs vary from 15 to 29 per cent. Larger differences are found between years than between models. With the exception of 1994, within specific years differences between models are at most 3 percentage points.
- For Finland and the aggregate group of immigrants with a citizenship of one of the other EU countries, models 1 (historical shares) and 2 (historical shares supplemented with stocks of foreigners) outperform the other models.
- Adding non-linear stocks together with internal migration variables (either net migration or outmigration) slightly improves the predictions for the aggregate group of non-EU nationals.
- Models 2 (historical shares and stocks of foreigners) and 3 (non-linear stocks of foreigners) are the best models for predicting the regional allocation of

Table IV. Prediction results (Relative Absolute Error), Sweden, NUTS 3, base year 1992 (stocks as per 31 Dec 1991)

	1993	1994	1995	1996	1997	1998	1999
Non-SE							
Model 1	0.18	0.29	**0.15**	**0.26**	0.19	0.22	0.27
Model 2	0.19	0.28	0.18	0.28	0.20	0.22	0.25
Model 3	0.18	0.26	0.17	0.28	0.19	0.22	0.26
Model 4.1	0.19	0.28	0.15	0.27	**0.18**	**0.21**	0.24
Model 4.2	0.18	0.25	0.16	0.27	0.18	0.21	**0.24**
Model 4.3	**0.16**	**0.22**	0.18	0.29	0.21	0.24	0.27
EU-other							
Model 1	0.11	0.18	0.15	0.10	**0.08**	**0.08**	**0.09**
Model 2	**0.10**	**0.17**	**0.14**	**0.09**	0.09	0.09	0.10
Model 3	0.22	0.26	0.25	0.20	0.19	0.15	0.16
Model 4.1	0.22	0.47	0.50	0.54	0.60	0.60	0.57
Model 4.2	0.22	0.27	0.25	0.20	0.19	0.14	0.16
Model 4.3	0.22	0.26	0.25	0.20	0.19	0.14	0.16
Non-EU							
Model 1	0.20	0.31	0.15	**0.27**	0.21	0.25	0.31
Model 2	0.20	0.29	0.20	0.32	0.24	0.26	0.30
Model 3	0.20	0.29	0.16	0.30	0.21	0.26	0.29
Model 4.1	**0.18**	**0.22**	0.21	0.36	0.27	0.31	0.34
Model 4.2	0.20	0.28	**0.15**	0.29	**0.19**	**0.24**	**0.26**
Model 4.3	0.19	0.25	0.18	0.31	0.22	0.27	0.31
Finland							
Model 1	**0.10**	0.12	**0.15**	0.14	0.13	0.14	**0.10**
Model 2	0.10	**0.12**	0.16	**0.13**	**0.12**	**0.14**	0.10
Model 3	0.28	0.30	0.24	0.26	0.30	0.26	0.25
Model 4.1	0.25	0.57	0.61	0.69	0.74	0.78	0.68
Model 4.2	0.28	0.30	0.24	0.26	0.30	0.26	0.25
Model 4.3	0.26	0.28	0.23	0.25	0.28	0.26	0.24
Iran							
Model 1	0.21	0.29	0.37	0.40	**0.32**	0.34	0.46
Model 2	**0.19**	**0.20**	0.26	0.31	0.37	**0.22**	**0.32**
Model 3	0.21	0.26	**0.25**	**0.29**	0.36	0.29	0.37
Model 4.1	0.26	0.43	0.47	0.53	0.50	0.48	0.57
Model 4.2	0.22	0.28	0.27	0.30	0.35	0.30	0.38
Model 4.3	0.21	0.27	0.26	0.29	0.36	0.29	0.36

Table continues on next page.

Table IV. Continued

	1993	1994	1995	1996	1997	1998	1999
Iraq							
Model 1	0.23	0.29	0.34	0.52	0.55	0.51	0.53
Model 2	0.28	0.35	0.35	0.40	0.39	0.36	0.29
Model 3	0.23	0.24	0.21	0.38	0.33	0.27	0.29
Model 4.1	0.30	0.41	0.44	0.65	0.59	0.56	0.53
Model 4.2	**0.19**	0.27	**0.20**	**0.33**	**0.30**	**0.25**	**0.24**
Model 4.3	0.20	**0.24**	0.23	0.39	0.36	0.30	0.32
South-America							
Model 1	0.26	0.18	0.12	0.16	0.26	0.23	0.27
Model 2	0.26	**0.18**	**0.09**	0.17	**0.24**	**0.19**	0.20
Model 3	0.26	0.22	0.18	0.16	0.27	0.23	0.27
Model 4.1	0.26	0.24	0.17	**0.14**	0.24	0.20	0.24
Model 4.2	0.20	0.23	0.20	0.21	0.30	0.20	**0.20**
Model 4.3	**0.19**	0.27	0.26	0.26	0.38	0.30	0.33

Bold numbers: the minimum RAE for the set of models in a given year for each immigrant group.

immigrants with an Iranian citizenship. The RAE of the best fitting models, however, is still rather high, varying from 19 to 32 per cent.

- For Iraq a significant improvement compared to models 1 and 2 is achieved when predictions are based on non-linear stocks supplemented with internal outmigration. For this group of immigrants too, the RAEs are relatively large (from 19 to 33 per cent).
- The results for the group of immigrants with a South-American citizenship are slightly different. For this group not one of the models in particular can be pointed as the best model tested. For all models, however, reasonable predictions are found (RAE varying from 9 to 24 per cent).
- Although in most cases – for different groups of immigrants as well as for different models – RAEs are higher for 1999 than for 1993, predictions do not systematically get worse as the projection horizon increases. Generally, RAEs show fluctuating highs and lows over the years. This volatility is due to the observed inflows, since the predicted inflows show a stable pattern over time. The results imply that there is no systematic bias in the time trend of the predictions.
- Predictions based on the year 1992 are slightly better than those based on 1989–1992. The difference is especially large for immigrants with an Iranian or Iraqi citizenship. This is not surprising though, as immigration of those groups was highly regulated until 1993. Therefore, regional immigration patterns before and after 1993 are probably influenced by rather different

Table V. Prediction results (Relative Absolute Error), Sweden, NUTS 3, base year 1989–92 (stocks as per 31 Dec 1988–91)

	1993	1994	1995	1996	1997	1998	1999
Non-SE							
Model 1	0.15	0.25	0.18	0.28	0.22	0.24	0.29
Model 2	0.15	0.24	0.21	0.32	0.24	0.26	**0.28**
Model 3	0.15	0.24	0.21	0.33	0.24	0.26	0.31
Model 4.1	0.16	0.25	0.20	0.32	0.23	0.26	0.30
Model 4.2	0.15	0.24	0.21	0.33	0.24	0.26	0.30
Model 4.3	**0.13**	**0.21**	0.23	0.34	0.25	0.28	0.32
Moving average	0.15	0.26	**0.17**	**0.27**	**0.21**	**0.23**	0.28
EU-other							
Model 1	0.14	0.19	0.16	0.14	0.17	0.15	0.16
Model 2	**0.13**	0.17	**0.13**	0.12	0.17	0.17	0.17
Model 3	0.25	0.29	0.26	0.23	0.21	0.17	0.18
Model 4.1	0.22	0.37	0.39	0.39	0.42	0.43	0.42
Model 4.2	0.24	0.29	0.25	0.23	0.21	0.17	0.18
Model 4.3	0.26	0.31	0.26	0.24	0.22	0.17	0.19
Moving average	0.14	**0.16**	0.15	**0.11**	**0.12**	**0.11**	**0.11**
Non-EU							
Model 1	0.16	0.27	**0.19**	0.32	**0.25**	0.28	0.34
Model 2	0.15	0.25	0.26	0.37	0.29	0.31	0.35
Model 3	0.17	0.27	0.22	0.36	0.26	0.31	0.35
Model 4.1	**0.13**	**0.20**	0.29	0.44	0.35	0.38	0.44
Model 4.2	0.16	0.26	0.22	0.36	0.26	0.30	0.35
Model 4.3	0.15	0.24	0.23	0.37	0.28	0.32	0.37
Moving average	0.16	0.28	0.19	**0.30**	0.25	**0.28**	**0.33**
Finland							
Model 1	0.14	0.13	0.18	0.14	0.12	0.18	0.14
Model 2	**0.14**	**0.13**	0.18	**0.13**	**0.11**	**0.16**	0.13
Model 3	0.25	0.28	0.23	0.25	0.28	0.28	0.23
Model 4.1	0.36	0.47	0.42	0.44	0.51	0.44	0.42
Model 4.2	0.29	0.31	0.25	0.27	0.31	0.27	0.23
Model 4.3	0.24	0.27	0.24	0.24	0.26	0.28	0.22
Moving average	0.14	0.13	**0.17**	0.14	0.12	0.17	**0.13**
Iran							
Model 1	0.36	0.42	0.52	0.55	0.42	0.46	0.60
Model 2	0.29	**0.29**	0.36	0.41	0.38	**0.29**	**0.40**
Model 3	**0.28**	0.34	0.37	0.40	**0.31**	0.38	0.48
Model 4.1	0.32	0.46	0.51	0.54	0.41	0.51	0.62
Model 4.2	0.29	0.35	0.37	0.41	0.32	0.39	0.49
Model 4.3	0.29	0.34	**0.36**	**0.39**	0.32	0.37	0.47
Moving average	0.36	0.38	0.46	0.46	0.39	0.42	0.55

Table continues on next page.

Table V. Continued

	1993	1994	1995	1996	1997	1998	1999
Iraq							
Model 1	0.44	0.46	0.57	0.74	0.77	0.72	0.75
Model 2	0.39	0.38	0.46	0.62	0.60	0.52	0.49
Model 3	**0.28**	**0.32**	0.39	**0.55**	0.53	0.45	0.48
Model 4.1	0.33	0.48	0.60	0.79	0.77	0.72	0.72
Model 4.2	0.28	0.33	0.39	0.56	0.53	0.46	0.48
Model 4.3	0.30	0.33	**0.39**	0.56	**0.51**	**0.43**	**0.45**
Moving average	0.44	0.43	0.50	0.66	0.74	0.68	0.69
South-America							
Model 1	**0.18**	0.24	0.27	0.31	0.40	0.36	0.44
Model 2	0.19	0.24	0.25	0.30	0.39	0.33	**0.36**
Model 3	0.26	0.28	0.30	0.32	0.40	0.35	0.41
Model 4.1	0.31	0.40	0.43	0.52	0.62	0.56	0.61
Model 4.2	0.24	0.30	0.32	0.36	0.45	0.36	0.40
Model 4.3	0.21	0.31	0.33	0.36	0.47	0.39	0.44
Moving average	0.18	**0.19**	**0.17**	**0.22**	**0.34**	**0.30**	0.36

Bold numbers: the minimum RAE for the set of models in a given year for each immigrant group.

phenomena. Due to the "Whole of Sweden" strategy, those migrant groups showed originally a more dispersed residential pattern compared to other migrant groups. After being placed in a sparsely populated area to start with, however, many immigrants moved to other areas. At the end of the 1980s, the Iranians' residential pattern became slightly more concentrated, which was contrary to the general trend in Sweden in which an increase of an immigrant group was accompanied by a more dispersed regional pattern (Borgegård et al., 1996).

- Looking at all models together, we may conclude that models 1 and 2 usually lead to the best predictions. Model 4.1 (including net internal migration) performs worse than the other models. Only for Iraq, including information on internal migration in addition to stocks of foreigners improves the prediction substantially.
- If we take into account moving averages over a 4-year period instead of a fixed average, the predictions slightly improve for most of the immigrant groups. Only for Iran and Iraq moving averages clearly outperform model 1 (RAEs decrease by 3 to 9 per cent points). On the other hand, for those immigrant groups, in none of the cases predictions based on moving averages turns out to be the best estimation. For Iranian and Iraqi immigrants, therefore, information on stocks, whether or not supplemented by information on internal migration, seems to be an essential component of regional allocation models in Sweden.

4. Conclusions

In this paper we explored whether it is feasible to improve the quality of subnational international migration assumptions by analysing the relationship between international migration flows and foreign population structures. A common method for projecting international migration at the regional level is allocating the projected national total immigration over the regions using regional shares. Different models for calculating immigration shares have been distinguished and empirically tested. These models are nested and may be characterized by the type of information they require: (1) models using historical information on the destination of migrants; (2) models using historical information and the regional distribution of stocks of foreigners; (3) models using non-linear stocks of foreigners; and (4) models using non-linear stocks of foreigners and information on aggregate internal migration (not specified by citizenship). Since the exact relationship between internal and external migration is not known, we estimated different variants of internal migration: net internal migration (model 4.1), internal outmigration (model 4.2) and the log of internal outmigration (model 4.3).

The models have been tested using data for Sweden for the following groups of immigrants: non-Swedish citizens, EU non-nationals, non-EU nationals, and nationals of Finland, Iran, Iraq, and South America. The models have been estimated using observed stocks of foreigners, immigration flows and internal migration patterns for 1992, as well as for the period 1989–1992. Subsequently, the estimated models have been used to predict the regional pattern of immigration for the period 1993–1999.

The results of the estimations and predictions show that in most cases more complicated models (models 3 and 4) do not improve significantly the results of the more simple models (models 1 and 2). Substantial improvements have only been found for immigrants with an Iranian or Iraqi citizenship. Moreover, although model 2 using historical shares supplemented with information on stocks of foreigners, often improves the prediction compared to model 1 using only historical shares, the improvement does not seem to make up for the loss of simplicity of the model. In order to be able to include information on stocks of foreigners in the models, some kind of prediction of regional stocks is needed, which is obviously far from being an easy task to perform.

More or less the same conclusion holds if we look at the difference between estimations and predictions based on observations for one base year compared to observations for a base period of a number of years (in this case a period of four years). Again, we could not conclude that taking into account observations for a couple of years rather than a single (most recent) year improves the results. Of course, here "the story behind the observations" may be more important than the observations themselves. For instance, if regulations concerning the regional allocation of refugees were different in the observation and the prediction period, regional shares based on the observation period may be poor predictors of the regional allocation of immigrants in the prediction period. Ideally, information on regulations and other immigration policies should be taken into account. On the other hand, in compiling regional population projections for a large number of countries, it is often not feasible to include all potentially important country-specific details. For such projections, calculating regional shares might be a reasonable solution. Based on the results of this study, however, we cannot favour one of the options above the other.

Overall, we may conclude that assumptions on the spatial distribution of immigration of foreigners can be improved by using the spatial distribution of stocks of foreign population as predictor. Improvements, however, are only minor and do not seem to compensate for the loss in simplicity of the models. This is especially true at the aggregate level of all immigrant groups. At this aggregate level, stocks of foreigners are rather static and therefore changes in stocks do not play an important part in changes in destination choices of immigrants. If we look at individual immigrant groups, on the other hand, information on stocks may improve the predictions. This seems to be especially the case for relatively new immigrant groups, like Iraqi immigrants in Sweden, where regional patterns are still on the move.

In situations where no regional breakdown of stocks of foreigners is available, in theory a model using regional shares based on the total population may be preferred over model 1, since the total population includes a crude indicator of stocks of foreigners. Predictions based on models with total population as proxy for stocks of foreigners, however, do not improve the results. On the contrary, in most of the cases models using historical shares only turned out to be the best predictors.

As in general the RAEs of the best predictions vary from about 10 to 30 per cent, predictions might be improved by adding additional variables to the models. For instance economic determinants, covering the economic business cycle or more structural developments like ageing of the labour force, may add to the explanation. Housing market variables too, may be of importance. Although taking into account those variables may further improve the models, the usefulness of such models is restricted to situations in which the external variables itself can be predicted with fair levels of accuracy, or to population scenarios in which the consequences of alternative economic developments are evaluated.

The results of the current study are somewhat at variance with those found in a comparable study for the Netherlands (Van der Gaag et al., 2001). In that study, the models were tested using data for the Netherlands for three groups of migrants (Turks, Moroccans and Surinamese) and the aggregate of the three groups. The models were estimated using observed stocks of foreigners, immigration flows and internal migration patterns for 1992, and the estimations were used to predict the regional pattern of immigration for 1995. There we found that generally the results *improve* with the complexity of the models. Inclusion of internal migration as predictor of international migration, however, seems to be justified only at the NUTS 2 level. At the NUTS 3 level, the addition of internal migration variables did not improve the results. Moreover, models for individual migrant groups perform marginally better than the model for the aggregate of the three groups. A complicating factor in comparing both studies is that they use different definitions of migrant groups. In the current study, country of citizenship was used, while in the study for the Netherlands, country of origin was used, defined as country of birth of the person himself, or, if the person was born in the Netherlands, the country of birth of his or her parents. Nevertheless, both studies show that for individual migrant groups, assumptions on the spatial distribution of immigration of foreigners could be improved by using the spatial distribution of stocks of foreigners as predictor. This may be especially important in regional projections of the population according to origin (see for instance Huisman and Van Wissen, 1998).

The main contribution of this paper is in the field of subnational population projections. Different relatively simple models have been presented requiring demographic data at the subnational level only. In many European countries, however, the necessary data for this modelling exercise are still difficult to obtain. Data availability is improving though (Rees and Kupiszewski, 1999), and similar studies for more countries are certainly called for.

As the level of geographical aggregation is important for the outcomes of the models, another line of research is also necessary. This can easily be explained with reference to the underlying theory of the attracting and facilitating roles of the migrant stock. Network relations that are relevant for attracting new migrants are most likely working at the very local scale. The linkages between stocks and flows should therefore ideally be studied at the very local level: zip codes, municipalities, city districts. Increasingly, these types of data are becoming available, for instance when population registers are linked with Geographical Information Systems. At this local level, also connections could be made with models of settlement concentration (e.g., Burnley, 1999).

Finally, these studies should also be supplemented by micro-level migrant studies (e.g., Esveldt et al., 1995). Although this may not be useful for regional forecasting purposes at first sight, it will enhance our understanding of the relationships between stocks and flows, and therefore improve the specification of subnational projection models.

Acknowledgements

This paper was written in the context of the research programme entitled "Towards a dynamic scenario model of economic determinants of European population dynamics", financially supported by the Netherlands Organisation for Scientific Research (NWO). Some parts of the paper are linked with the study "Regional International Migration and Foreign Population within the European Union", a feasibility study commissioned and funded by the European Commission, DG-XVI. An earlier version of this paper was presented at the BSPS/NVD Conference

"Population Studies in Britain and the Netherlands", Utrecht, 2000. We would like to thank two anonymous referees for their very useful comments on an earlier draft.

References

Böcker, A., 1994. 'Chain migration over legally closed borders: Settled immigrants as bridgeheads and gatekeepers'. *Netherlands Journal of Social Sciences* 30(2): 87–106.

Borgegård, L., J. Håkansson and D. K. Müller, 1996. 'The changing residential patterns of immigrants – the case of Sweden 1973–1992'. *Yearbook of Population Research in Finland* 33: 173–183.

Boyd, M., 1989. 'Family and personal networks in international migration: Recent developments and new agendas'. *International Migration Review* 23(3): 638–670.

Bucher, H., 1996. 'Migration and impact on internal distribution and redistribution of population', in M. Frey and U. Mammey (eds), *Impact of Migration in the Receiving Countries, Germany*, pp. 45–47. International Organization for Migration, Geneva.

Burnley, I., 1999. 'Levels of immigrant residential concentration in Sydney and their relationship with disadvantage'. *Urban Studies* 36(8): 1295–1316.

Coleman, D., 1994. 'International migrants in Europe: Adjustment and integration processes and policies', in M. Macura and D. Coleman (eds), *International Migration: Regional Processes and Responses*, pp. 41–76. United Nations, New York.

De Jong, A. and Visser, H., 1997. 'International migration scenarios for the European Economic Area', Eurostat Working Paper E4/1997-6, Eurostat, Luxembourg.

Edmonston, B. and Passel, J. S., 1992. 'Immigration and immigrant generations in population projections'. *International Journal of Forecasting* 8: 459–476.

Esveldt, I., Kulu-Glasgow, I., Schoorl, J. and Van Solinge, H., 1995. *Migratiemotieven, migratienetwerken en partnerkeuze van Turken en Marokkanen in Nederland* (Migration motives, migration networks and partner choice of Turks and Moroccans in the Netherlands), NIDI report 43, NIDI, The Hague.

Fawcett, J. T., 1989. 'Networks, linkages, and migration systems'. *International Migration Review* 23(3): 671–680.

Frey, W. H. and Liaw, K. L., 1998. 'The impact of recent immigration on population redistribution within the United States', in J. P. Smith and B. Edmonston (eds), *The Immigration Debate: Studies of the Economic, Demographic and Fiscal Effects of Immigration*, pp. 388–448. Washington, DC: National Academy Press.

Gurak, D. T. and Caces, F., 1992. 'Migration networks and the shaping of migration systems', in M. M. Kritz, L. L. Lim and H. Zlotnik (eds), *International Migration Systems: A Global Approach*, pp. 150–176. Oxford: Clarendon Press.

Hugo, G. J., 1981. 'Village-community ties, village norms, and ethnic and social networks: a review of evidence from the third world', in G. F. de Jong and R. W. Gardner (eds), *Migration Decision Making: Multidisciplinary Approaches to Microlevel Studies in Developed and Developing Countries*, pp. 186–224. New York: Pergamon Press.

Huisman, C. and Van Wissen, L., 1998. *Regionale Allochtonen prognose 1996–2016* (Regional forecast of foreign population 1996–2016), NIDI report 54, NIDI, The Hague.

Leering, T. J. and Den Otter, H. J., 1992. 'Buitenlandse migratie en regionale ontwikkelingen op COROP-niveau', INRO-TNO publication 92/NW/065, TNO, Delft.

Massey, D. S., Arango, J., Hugo, G., Kouaouci, A., Pellegrino, A. and Taylor, J. E., 1993. 'Theories of international migration: a review and appraisal'. *Population and Development Review* 19(3): 431–466.

Münz, R., 1996. 'A continent of migration: European mass migration in the twentieth century'. *New Community* 22(2): 201–226.

Rees, P. and Kupiszewski, M., 1999. 'Internal migration: what data are available in Europe?'. *Journal of Official Statistics* 15(4): 551–586.

Rees, P., Kupiszewski, M., Eyre, H., Wilson, T. and Durham, H., 2001. 'The evaluation of regional population projections for the European Union', Eurostat Working Paper 3/2001/E/no.9, Eurostat, Luxembourg.

Stillwell, J., Eyre, H. and Rees, P., 1999. *Regional International Migration and Interregional migration within the European Union: A Feasibility Study*. Study prepared for the Commission of the European Communities, DGXVI and Eurostat, University of Leeds, Leeds.

Van der Gaag, N. and Van Wissen, L., 1999. *Analysis and Forecasting of International Migration by Major Groups (Part II)*. Eurostat Working Paper 3/1999/E/no.9, Eurostat, Luxembourg.

Van der Gaag, N. and Van Wissen, L., 2001. 'Determinants of the subnational distribution of immigration'. *Journal of Economic and Social Geography* 92(1): 27–41.

Van der Gaag, N., Van Wissen, L., Van Imhoff, E. and Huisman, C., 1999. 'National and regional population trends in the European Union', Eurostat Working Paper 3/1999/E/no.8, Eurostat, Luxembourg.

Van der Gaag, N., Van Wissen, L., Salt, J., Lynas, Z. and Clarke, J., 2001. 'Regional international migration and foreign population within the EU; a feasibility study', Eurostat Working Paper 3/2001/E/no.5, Eurostat, Luxembourg.

Van Huis, L.T. and Nicolaas, H., 1999. 'Verhuizingen in de vier grote steden, 1997' (Big city migration, 1997). *Maandstatistiek van de bevolking* 99(5): 9–17.

Van Imhoff, E., Van Wissen, L. and Spiess, K., 1994. 'Regional population projections in the countries of the European Economic Area', NIDI/CBGS publications no. 31, Swets and Zeitlinger, Amsterdam/Lisse.

Van Wissen, L. and Visser, H., 1998. 'An explanatory analysis of international migration flows within the European Economic Area', in C. Gorter, P. Nijkamp and J. Poot (eds), *Crossing Borders: Regional and Urban Perspectives on International Migration*. Ashgate: Aldershot.

Waldorf, B., 1996. 'The internal dynamic of international migration systems'. *Environment and Planning A* 28: 631–650.

White, P., 1993. 'Immigrants and the social geography of European cities', in R. King (ed), *Mass migration in Europe, the Legacy and the Future.* London: Belhaven Press.

Wilpert, C., 1992. 'The use of social networks in Turkish migration to Germany', in M. M. Kritz, L. L. Lim and H. Zlotnik (eds), *International Migration Systems; A Global Approach.* Oxford: Clarendon Press.