



Economic Commission for Europe

Conference of European Statisticians

Sixty-second plenary session

Paris, 9-11 April 2014

Item 4 of the provisional agenda

Migration statistics**Forecasting migration in official population projections using
an econometric model****Prepared by Statistics Norway***Summary*

This paper introduces an approach to compiling immigration projections using an econometric model based on standard migration theories. Long-term population developments and the related factors are of interest to decision makers, scientists and many other users. Although substantial research has been made in quantifying determinants of international migration, most official population projections do not include these determinants in an explicit migration model. The main variables used in the immigration projections include income level, unemployment and population size in the receiving and sending countries, as well as previous immigration from the sending countries to the receiving country, Norway.

The paper is presented for discussion to the Conference of European Statisticians seminar on migration statistics.

I. Introduction

1. Long-term population developments and the related factors are of interest to decision makers, scientists and many other users. In Norway, migration constitutes an important part of population change. Based on surveying the practice of international organizations and national statistical offices in migration forecasting, large differences exist between the agencies – regarding practical and technical sides of the projections, and also in the factors included when explaining migration changes. Most forecasts rely on demographic and economic reasoning. However, they are not based on models, at least not formalized ones, which can be assessed by an outside observer and evaluated against alternatives. The paper emphasizes the importance of formal models where all assumptions are explicitly stated regarding model specification and paths of exogenous variables that are expected to influence the demographic component of interest. Following such an approach exposes to criticism which has the potential of yielding a better model specification and more realistic trajectories of different components of the population in the next round of calculations.

II. Model framework

2. The paper introduces the formal model used for forecasting gross immigration to Norway until 2100. Our modelling framework is based on standard micro theory on factors that are important for a decision to migrate. The focus in this literature is on expected income differences adjusted for migration costs. In addition political factors affecting the decision to migrate may be important. We distinguish between immigration to Norway from three geographical areas, *Western countries*, *Eastern EU members* and *Rest of the world*¹. Separate models, where the log-transformed migration rate (migration to Norway divided by the population in each area) is the endogenous variable, are estimated for these three areas. The explanatory variables as well as the dynamic specification differ somewhat between the three areas.

3. The results of the estimation are shown in Table 1. In all areas, and particularly for *Eastern EU members*, the ratio between PPP-adjusted gross domestic product per capita in Norway and the corresponding measure in the sending area is a significant explanatory variable. For *Western countries* and *Eastern EU members* both the unemployment rate in Norway and in the sending area plays a role. However, only in the latter the unemployment level enters significantly. For *Rest of the world* we have no reliable time series for the unemployment rate, so for this area only the (change) in the Norwegian unemployment rate enters the econometric specification. In contrast to the other areas, a “bridge head” variable turns out to be significant for *Rest of the world*. This variable was implemented as the stock of immigrated people from the area scaled by the Norwegian population. The lagged endogenous variable is included for all the three areas. For *Eastern EU members* and *Rest of the world* we also include some dummy variables picking up, among other things, institutional changes.

¹ “Western countries” includes Western Europe, USA, Canada, Australia and New Zealand; “Eastern EU-members” are the 10 Eastern European countries included in the EU in 2004 or 2007; and “Rest of the world” includes all other countries, including all of Africa, Asia and Latin America. The three areas are characterized by different reasons for migration: Labour among the *Eastern EU* immigrants, flight and family among migrants from *Rest of the world*, whereas migration from *Western countries* is dominated by labour and family immigration.

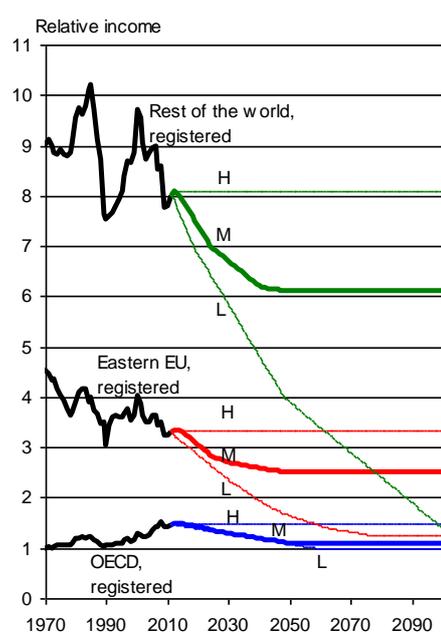
III. Forecasts of exogenous variables

4. The empirical models are used for forecasting gross immigration from the three areas to Norway until 2100. To achieve this, explicit assumptions with respect to trajectories for the exogenous variables have to be made. We consider three conditional forecasts – high, medium and low. Different assumptions with respect to development in relative income and treatment of forecasting uncertainty are the reasons why the three forecasts differ.

5. The high (H), medium (M) and low (L) forecasts for income differences between Norway and the three country groups are shown in Figure 1. They are based on quite different scenarios for the Norwegian and the global economy. In the medium trajectory, the Norwegian purchasing power parity (PPP) adjusted gross domestic product (GDP) per capita, which is now 40 percent higher than the OECD level, is expected to decrease to 10 per cent above the OECD level from 2050. This is due to an expected reduction of the Norwegian petroleum income.

Figure 1

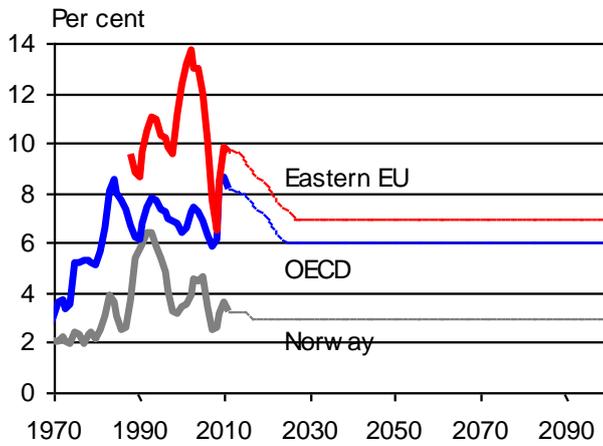
Annual relative income per person in Norway compared to the three country groups. Registered 1970-2011 and assumed 2012-2100 in three alternatives



6. The forecasts are aligned with the United Nations projections for the development of the population in the three sending areas.

7. The assumptions for the development of the unemployment rates (shown in Figure 2) are also the same for all the three conditional forecasts.

Figure 2
Unemployment rate in Norway, OECD and Eastern EU. Registered 1970-2011 and assumed 2012-2100



8. Using these forecasts, three trajectories are made for future immigration to Norway from each of the three country groups. In the low (L) alternative we have subtracted an estimate of the model's standard error from the point forecast to indicate forecasting uncertainty stemming from the error terms. The same is added to the high (H) alternative.

9. The results for each country group and the total immigration to Norway can be seen in Figures 3-6:

Figure 3
Annual immigration from Western countries. Registered 1970-2011 and projected 2012-2100 in three alternatives

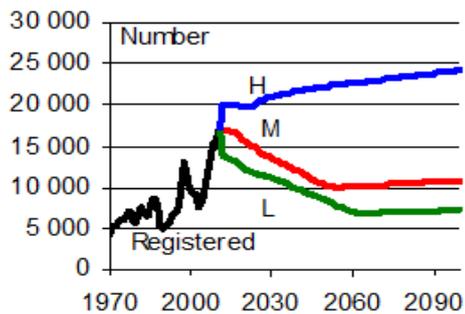


Figure 4
Annual immigration from Eastern EU members. Registered 1970-2011 and projected 2012-2100 in three alternatives

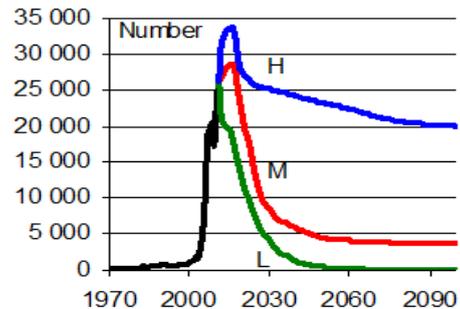


Figure 5
Annual immigration from the rest of the world. Registered 1970-2011 and projected 2012-2100 in three alternatives

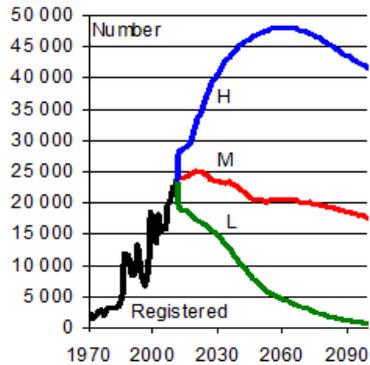
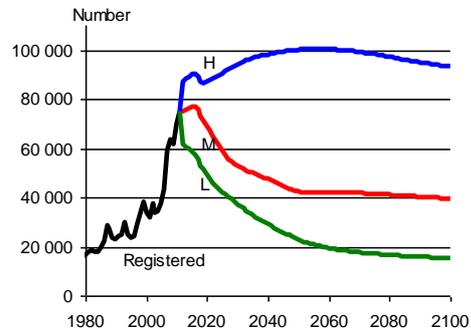


Figure 6
Total annual immigration to Norway (including a constant immigration of 7 440 persons of Norwegian origin). Registered 1970-2011 and projected 2012-2100 in three alternatives



10. The forecasts suggest that total immigration is likely to stay high at least for another decade or so. However, the different forecast alternatives indicate that immigration is highly uncertain, and even if we condition on a specific path of the explanatory variables the prediction intervals are wide.

11. The treatment of forecast uncertainty is an important issue in long term forecasts. If the model specification is correct one may distinguish between three different sources for forecast uncertainty, i.e., (i) uncertainty brought about by the error term in the econometric models underlying the forecasts, (ii) uncertainty in parameter estimates and (iii) uncertainty related to the exogenous variables. In the current paper we have ignored the second one and taken the last type into account by considering different trajectories for the income ratio variables. However, the estimated prediction intervals using the existing models indicate that even our conditional forecasts are highly uncertain, as shown in Figure 7.

Figure 7
Projected total immigration in the Medium alternative, including estimated prediction intervals (+/- one standard deviation)

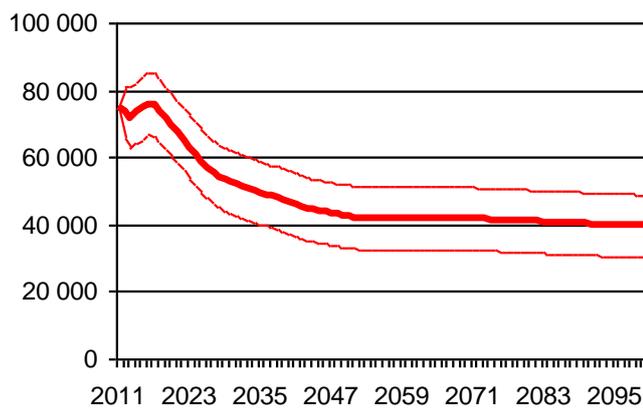


Table 1
 Estimation results for $\log(M/P)_t$ for the three areas and regression diagnostics²

Explanatory variables	Western countries		Eastern EU-members		Rest of the world	
	Estimate	t-value	Estimate	t-value	Estimate	t-value
Constant	0.815	5.58	-0.713	-1.02	-0.741	-0.78
$\log(M/P)_{t-1}$	0.627	9.25	0.632	7.12	0.486	4.06
$\log(Y_{NOR,t}/Y_t)$			1.228	2.52		
$\log(Y_{NOR,t-1}/Y_{t-1})$	0.773	4.48				
$\log(Y_{NOR,t-2}/Y_{t-2})$			0.898	2.00		
$\log(Y_{NOR,t-3}/Y_{t-3})$					1.255	2.78
$\log(IS_{t-1}/POP_{NOR,t-1})$					0.343	4.05
$\Delta U_{NOR,t}$			-0.208	-3.74		
$\Delta U_{NOR,t-1}$					-0.096	-1.71
$\Delta^2 U_{NOR,t}$	-0.099	-7.34				
$U_{NOR,t-3}$			-0.132	-4.27		
ΔU_{t-1}	0.045	2.19				
U_{t-1}			0.031	1.61		
DUM6704			-0.924	-6.33		
DUM07			0.390	2.10		
DUM77					-0.266	-2.39
DUM87					0.605	3.19
DUM93					0.596	3.13
R ²	0.934		0.997		0.947	
Estimation sample	1970–2011		1990–2011		1970–2011	
Regression diagnostics						
AR 1-2 test (residual autocorrelation)	F(2,35)=1.254, pvalue = 0.298		F(2,11)=2.886, pvalue = 0.098		F(2,32)=0.728, pvalue = 0.320	
ARCH 1-1 test (residual heteroskedasticity)	F(1,40)=0.397, pvalue = 0.532		F(1,20)=1.3551, pvalue = 0.258		F(1,40)=0.049, pvalue = 0.826	
Normality test	$\chi^2(2)=0.202$, pvalue = 0.904		$\chi^2(2)=0.918$, pvalue = 0.632		$\chi^2(2)=4.156$, pvalue = 0.125	
Hetero test (residual heteroskedasticity)	F(14,27)=1.406, pvalue = 0.217		F(15,6)=2.033, pvalue = 0.195		F(15,24)=0.719, pvalue = 0.743	
RESET23 test (Functional misspecification)	F(2,35)=2.058, pvalue = 0.143		F(2,11)=3.376, pvalue = 0.632		F(2,32)=1.509, pvalue = 0.236	

² The left hand side variable, $\log(M/P)$, is the log of the migration rate. Y_{NOR} and Y denote, respectively, PPP-adjusted GDP per capita for Norway and the sending area. IS denotes the immigration stock of the sending area and POP_{NOR} denotes the Norwegian population. U_{NOR} and U denote, respectively, the unemployment rate in Norway and in the sending area. The variables $DUM6704$, $DUM07$, $DUM77$, $DUM87$ and $DUM93$ are dummy variables picking up the effects of policy changes and special events.