Paweł Strzelecki[*]
Instytut Statystyki i Demografii, Szkoła Główna Handlowa
Institute of Statistics and Demography
Warsaw School of Economics

THE MULTI-STATE PROJECTION OF POLAND'S POPULATION BY EDUCATIONAL ATTAINMENT FOR THE YEARS 2003–2030

INTRODUCTION

Education is one of the most important factors to be accounted for in the analyses of various social and economic processes. It has far reaching significance, both at the individual and societal levels. At the individual level better education means, among others, improved quality of life, more job opportunities, higher expected wages, openness to innovation, further learning ability, etc. At the societal level population composition by education is to be identified as a key factor in development and economic performance of countries (Mankiw et al., 1992). Increasing skills of the population complement and enhance all other development factors and explain a significant part of the cross-country differences in economic growth (Cohen and Soto, 2001; Florczak, 2008).

Moreover, the educational attainment of the population has overwhelming implications for demographic processes. The relationships between education and fertility are complex and have been widely studied. The postponement of maternity is to a large extent linked to a longer time period spent in education. Economic theories of fertility refer to the opportunity costs of children increasing with higher education levels, which also contribute to lower fertility (Becker et al., 1990). Relationships between education level and fertility decisions in Poland have been confirmed as well (Jóźwiak et al., 2000; Kotowska et al., 2007).

Furthermore, many international studies have shown mortality differentials by education level (e.g., Huisman and Kunst, 2004; Shkolnikov et al., 2006). Mortality

[*] E-mail: pstrzel1@sgh.waw.pl. The author was a participant of the Young Scientist Summer Program 2005 in IIASA, during which he acquired valuable experience in using multi-state projection models. All possible mistakes are the only responsibility of the author and not the institutions with which he cooperated.

differentials by education have been also noticed in Poland (Wróblewska, 2006). Although the relationship between the number of years spent at school and life expectancy is not clear-cut, the evidence of that relationship is compelling. Theories that attempt to explain the positive influence of years of schooling on life expectancy usually give prominence to factors such as: differences in income, access to better health care, healthier lifestyle, etc.

The educational attainment could be also used as an additional variable in population projections. This study follows the literature on the application of the multi-state projection methodology to the projections of population by the achieved level of education. Most of the publications in this field refer to the papers of the demographic department of the International Institute for Applied System Analysis (IIASA) in Vienna. In the publications of IIASA educational attainment was used on one hand as a powerful predictive factor of demographic changes (Goujon and Wils, 1996; Lutz et al., 1998). On the other hand, the structure of population by education level, was frequently used to measure human capital (Lutz and Goujon, 2001).

The analyses presented aim to answer two questions. The first question is: how would the structure of the population of Poland change, with regard to educational attainment, age and sex, under simple assumptions about basic components of population dynamics. It is assumed that both mortality and fertility age-sex--education patterns, as well as transition probabilities of changing educational attainment would stay constant in the future and there would be no migration. The second research question is: what are the contributions of the changing population structure by educational attainment to the changes in predicted aggregate measures of fertility and mortality, under the assumption that the differences in mortality and fertility by education would be constant in the future.

The article starts with a discussion of detailed assumptions of the multi-state population projection model applied to the projection of the population of Poland by age, sex and education level, for the years 2002–2030. The word "projection" is used here to describe a method of predicting the population size and its composition under certain assumptions[1]. The next section presents the projection results for the years 2003–2007, for which empirical observations exist and can be compared with. The results can be used in various socio-economic analyses. The comparison shows that the constant-parameter assumptions lead to the results that are relatively similar to the observed empirical data. This means that for such a relatively short period of time these projections performed relatively well. This fact suggests that despite the simplicity of the assumptions, the projection results can be used as a forecast, within a short forecast horizon.

---

[1] We distinguish between "projection" and "forecast" proposed in UN (1958) (after Preston et al., 2001). According to this definition: "projections are the calculations which shows the future development of a population when certain assumptions are made about the future cause of fertility, mortality, and migration. A population forecast is a projection in which the assumptions are considered to yield a realistic picture of the probable future development of population" (Preston et al., 2001: 117).

Subsequently, the projection results are used to simulate possible future changes in the aggregate measures of fertility and mortality which may be due to the shifts in the population structure by educational attainment. The influence of the changes in the education structure on mortality in selected countries of Eastern Europe was recently studied by Shkolnikov et al. (2006). One of the possible applications of our projection results is a similar analysis for the future population of Poland. Such an analysis can shed some light on the likely future growth in life expectancy due to improvement in education.

The detailed analysis of the Poland's population has been carried out for single-year age groups and seven levels of education. It adds an additional dimension to the official population forecast for Poland with base year 2003 (Bolesławski, 2004), and to population projections by Matysiak and Nowok (2007), both taking into account the results of the last Population Census from 2002. This projection also continues the idea of constructing multi-state projection models for Poland. According to the author's knowledge, the multi-state projection models for Poland were used in population projections that took into account migration flows between countries and regions (Kupiszewska and Kupiszewski, 2005) and the structure of population by marital status (Kotowska, 1994). The multi-state model was also used to regional analysis of fertility (Paradysz, 1981) and to some extent in multidimensional studies of population mentioned in literature review made by Paradysz (1988).

THE MULTI-STATE MODEL

In this study we apply a multi-state projection model to prepare a population projection, which includes educational attainment. Multi-state models has been used in population projections since 1970s. Originally, the new dimensions added to the population projection were geographical units (regions). The multi-regional projection model was initially developed by Rogers (1975) as a synthesis of multidimensional expansion of life-tables and the cohort-component projection model. Subsequently, the multi-regional method was extended and appeared as a multi-state model. The short introduction to the initial idea and the selected applications can be found in Rogers (1980). Since the beginning of the 1980s the multi-state models have been applied and extended by many authors (i.e. Willekens, Drewe, 1984). The model used in our study was formulated by van Imhoff and Keilman (1991) and implemented in LIPRO 4.0 software. This software enables not only all calculations of predicted population variables but also the preparation of assumptions by estimation of the matrices of intensities from the data on one-year transitions[2].

The state space of the multi-state model (van Imhoff and Keilman, 1991) is defined by the primary demographic characteristics of the individuals that are taken into consideration, i.e. education. The state vector, which describes the population

[2] The software was developed in the Netherlands Interdisciplinary Demographic Institute (NIDI). For more information about implementation of the multi-state method in LIPRO 4.0 see van Imhoff (1994).

at the given time point, consists of numbers of individuals at that time, broken down by age, sex and educational attainment. The population changes over time are due to events experienced by individuals, which can be divided into internal and external ones. The former refer to the changes in education, i.e. transitions of individuals between different education levels, while the latter means entries to the population (births, immigration) and exits from the population (deaths, emigration). It should be also mentioned that the transitions between the particular states describing education are only in one direction: from less to more advanced levels, which simplifies the analysis. Moreover, although the population under consideration is not closed, because of the "zero migration" assumption only endogenous entries (births) and exists (deaths) are distinguished as external events.

DATA SOURCES AND ASSUMPTIONS

The projection accuracy of the model used in this paper is a compromise between an attempt to construct a very detailed model and the limitations of available data. To predict the population year by year, from 2002 to 2030, a relatively high detail of age profiles was required. Finally, seven education groups, consistent with the ISCED classification (UNESCO 2006, MEN 2008) were included in the projection for single-year-age groups (Table 1). This means that the population is broken into 100 x 2 x 7 = 1400 cells, taking into account age (100 groups), sex (2 groups) and completed education level (7 groups).

Table 1. Completed levels of education used in the projection model and ISCED classification
Tabela 1. Poziomy wykształcenia w projekcji i według klasyfikacji ISCED

Educational attainment in the projection *Poziom wykształcenia w projekcji*	ISCED classification *Klasyfikacja ISCED*
Tertiary Wyższe	ISCED 5, 6
Post-secondary Policealne	ISCED 4
Secondary vocational Średnie zawodowe	ISCED 3A
Secondary general Średnie ogólne	ISCED 3A
Basic vocational Zasadnicze zawodowe	ISCED 3C
Primary Podstawowe	ISCED 1, 2
Less than primary and no education *Niższe niż podstawowe i brak wykształcenia*	-

Source: Own elaboration on the basis of UNESCO (2006) and MEN (2008).
Źródło: Opracowanie własne na podstawie UNESCO (2006) i MEN (2008).

A person attains a certain level of education when he/she graduates from a school or an university. This means, for example, that university students are not considered as persons with tertiary education, until they graduate with Bachelor's or Master's degrees. Before that, they are classified as persons with secondary or post-secondary education.

The Population Census 2002 was the most important source of data to build the projection model. The structure of the initial population by sex, single-year age groups and different educational attainment was taken directly from the census (CSO 2003d). The Karup-King interpolation formula (Shyrock and Sieger 1976) was applied to obtain one-year age groups, when the cross-sectional data were available in 5-year age groups only. It should be mentioned that despite the fact that the census data covers the whole population, one should take into account possible inaccuracies. First question concerns the applied definition of the population, which in the case of the official annual population statistics for Poland suffers from the lack of clarity. In fact, the concept used in official statistics is something between *de facto* and *de jure* population and could not, in all conscience, be used in population projections (Kupiszewski et al., 2003). However, the lack of other, appropriately detailed data, as well as the requirement of comparability with other official education statistics, forced the author to use the census data without any corrections, as the initial population[3]. Another source of inaccuracy could be the percentage of answers: "unknown educational attainment", amounting to about 2% of the total population. In our projection this problem was solved by spreading this 2% group proportionally between different education groups. Finally, it should also be remembered, that answers given in the census and in the Labour Force Survey (LFS) were not verified with documents confirming the formal education level.

The Labour Force Survey (LFS) is another source of data on educational attainment. About 0.2% of the population of Poland is quarterly covered by this survey. The data on the education structure available from this survey is to some extent biased in comparison to the total population, because the LFS data does not cover the population living in collective households, such as students' dormitories, prisons or social welfare homes. In general, this may lead to the underestimation of the share of young people with secondary and higher education, who often live in dormitories. The availability of micro-level data from the LFS was a strong advantage, because different types of cross-sectional data were required for the projection and its verification. What is more important, the LFS sample data allow to analyze single-year changes of person's educational attainment, which provided very valuable information for the calibration of flows between different education levels. Despite the fact that the LFS is one of the biggest regular surveys in Poland, the main limitation of the LFS database, in the context of this analysis, was the

[3] The definition problems could also cause errors in the estimation of fertility and mortality rates calculated by CSO, but there are no other estimates of these values for the latest data. The influence of "invisible" migration (not registered by CSO) in the 1980s on the population size and the age composition and estimates of fertility and mortality rates were presented by Sakson (2002).

size of the sample, which did not allow a direct and reliable calculation of flows between different education levels. That is why the preparation of parameters for the multi-state projection required many efforts and usage of different techniques and data sources, which are described below.

We assumed morality rates to be constant at the level observed in the year 2002, and therefore the Polish life tables for 2002 (CSO 2003c) were the main source of data. However, the differences between age-specific mortality rates for groups with different education levels were introduced in the projection. The information on relative differences in mortality by education was taken from the publication of Wróblewska (2006) and confronted with data available on Polish CSO web pages (CSO 2007b), as well as with results obtained for other countries (Lutz et al. 1998, Huisman and Kunst 2004). Life tables for groups of people with different education status were calculated under the assumption that age-specific probability of death for a person from the general population is a weighted average of death probabilities of persons from different education groups, as given in (1):

$$q(x,s) = \sum_{edu} q^{ref}(x,s) \cdot \omega(x,s,edu) \cdot r(x,s,edu). \qquad (1)$$

In the above formula, x denotes age, s – *sex* and *edu* – the level of education (e.g.: 1 – "Tertiary" 2 – "Postsecondary", etc.), $q^{ref}(x,s)$ stands for the probability of death within the age group x, for sex s. Given the reference education level (here, the "Tertiary education"), $r(x,s,edu)$ is the relative probability of death of a person, aged x and of sex s, in the group with educational attainment *edu,* in comparison with the reference group. By definition, for the reference group $r(x,s,ref)=1$. The expression $\omega(x,s,edu)$ represents the share of the group, with selected age, sex and education level, in the population of the same age and sex.

The formula for the probability of death of a person from the reference education group could be derived from (1) as:

$$q^{ref}(x,s) = \frac{q(x,s)}{\sum_{edu} \omega(x,s,edu) \cdot r(x,s,edu)}. \qquad (2)$$

Under this assumption, the probabilities of death by sex, age and education were calculated separately for men and women, using information on: the probabilities of death in the total population, relative differences between different educational groups, and the shares of these groups in the population. Formula (3) presents the method used to calculate the probability of death for an individual from a specific education level k.

$$q(x,s,k) = q^{ref}(x,s) \cdot r(x,s,k). \qquad (3)$$

In 2002, the life expectancy at birth amounted to 70.4 years for men and 78.8 years for women. In general, the differences in mortality rates by age were larger

among men, especially at younger ages. Life expectancy at birth for different education groups varied from 65.8 to 79.8 years for men and from 76.6 to 86.1 years for women. The highest value of life expectancy was observed for those men and women, who achieved tertiary education before the age 25. Persons with primary or uncompleted primary education during the whole life-course have the lowest life expectancy (Table 2).

Table 2. Comparison of the relative mortality rates and life expectancy for different population groups by education

Tabela 2. Porównanie relatywnych współczynników zgonów i oczekiwanej długości trwania życia grup ludności o różnym poziomie wykształcenia

Sex / Education level *Płeć / Poziom wykształcenia*	Relative mortality rates by age in comparison to the reference group *Relatywne współczynniki zgonów według wieku w porównaniu do grupy referencyjnej*						e(0)
	20–29	30–39	40–49	50–59	60–69	70+	
Males (total) *Mężczyźni (ogółem)*							70.4
Tertiary & postsecondary* *Wyższe i policealne**	1.00	1.00	1.00	1.00	1.00	1.00	79.8
Secondary general & vocational *Średnie ogólne i zawodowe*	1.83	2.20	2.03	1.79	1.59	1.36	74.9
Basic vocational *Zasadnicze zawodowe*	3.55	4.32	3.90	3.19	2.90	2.33	67.9
Less than primary & primary *Mniej niż podstawowe i podstawowe*	5.68	8.89	6.64	3.65	2.53	2.19	65.8
Females (total) *Kobiety (ogółem)*							78.8
Tertiary & postsecondary* *Wyższe i policealne**	1.00	1.00	1.00	1.00	1.00	1.00	86.1
Secondary general & vocational *Średnie ogólne i zawodowe*	1.60	1.38	2.56	2.08	1.49	1.53	81.1
Basic vocational *Zasadnicze zawodowe*	2.10	2.61	3.32	2.61	3.47	2.29	77.0
Less than primary & primary *Mniej niż podstawowe i podstawowe*	6.25	6.16	4.48	2.61	2.58	2.36	76.6

* Reference group – Tertiary and postsecondary education.

* *Grupa referencyjna – Wykształcenie wyższe i policealne.*

Source: Own calculations after Wróblewska (2006).

Źródło: Obliczenia własne na podstawie pracy Wróblewskiej (2006).

The assumptions on fertility are based on the distribution of age-specific fertility rates (ASFRs) for the 2002 population. Additionally, differences between the fertility rates for women with different educational attainment were taken into account. Calculation of fertility rates by education was similar to the approach used for the probabilities of death. Under the assumption on the age-specific fertility

rate of women at a certain age, $ASFR(x)$, being a weighted average of age-specific fertility rates of women with different educational attainments, the formula used in calculation is given below:

$$ASFR(x) = \sum_{edu} ASFR^{ref}(x) \cdot r(x,edu) \cdot \omega(x,edu) \cdot \tag{4}$$

In (4), $ASFR^{ref}(x)$ is the age specific fertility rate of women in the reference group (in this paper the reference group for fertility was "primary and uncompleted primary")[4], $r(x,edu)$ denotes relative fertility rates in comparison to $ASFR^{ref}(x)$, and $\omega(x,edu)$ is the share of the education group *edu* in the whole population of women aged *x*. The-age specific fertility in the reference group can be thus derived from (4) as:

$$ASFR^{ref}(x) = \frac{ASFR(x)}{\sum_{edu} r(x,edu) \cdot \omega(x,edu)}. \tag{5}$$

The age-specific fertility rate for the selected education level *k* may be obtained as:

$$ASFR(x,k) = ASFR^{ref}(x) \cdot r(x,k). \tag{6}$$

The relative differences between the age-specific fertility rates for women with different education levels were calculated from the data on the number of births and the data on the female population in the year 2002 (Table 3).

The highest total fertility rate (TFR) was observed in the group of women with basic vocational education. That was mainly the result of high fertility in the 15–30 years age group. The acquired education level clearly influenced the age of the highest fertility in comparison with other groups. For women with tertiary education, the highest fertility was noted in the 30–34 age group, while for those with secondary education it was observed between 25 and 29 years of age. The calculation of fertility rates based on the period data only does not take into account recent changes in cohort fertility. It is an important limitation of the approach used here.

The assumptions regarding the probabilities of transitions between different education levels are most important for modeling changes in the population composition by educational attainment. The probabilities used in this projection are the results of an analysis of data from the 2002 census and of the transition probabilities estimated on the 2002–2003 LFS data. Firstly, transitions were restricted only to those considered as possible[5]. Secondly, an evaluation of education statistics

[4] This category is a sum of two subcategories: "no education or less than primary" and "primary education" which were taken into account in the projection.

[5] If the transitions between all 7 states would be possible during the whole life period (101 age categories and 2 sexes), the number of transition parameters in the model would reach 9,800.

Table 3. Comparison of age-specific relative fertility rates and TFRs for different population groups by educational attainment, 2002

Tabela 3. Porównanie względnych współczynników płodności według wieku oraz TFR w grupach kobiet o różnym poziomie wykształcenia, rok 2002

Educational attainment *Poziom wykształcenia*	Relative fertility rates by age groups *Względne współczynniki płodności według grup wieku*							TFR by educational attainment *TFR według poziomu wykształcenia*
	15–19	20–24	25–29	30–34	35–39	40–44	45–49	
Tertiary & postsecondary *Wyższe i policealne*	0.00	0.35	0.81	1.07	0.76	0.55	0.69	0.99
Secondary general & vocational *Średnie ogólne i zawodowe*	1.83	0.45	1.02	0.90	0.67	0.64	0.86	1.13
Basic vocational *Zasadnicze zawodowe*	7.61	1.08	1.10	0.96	0.85	0.81	1.06	1.65
Less than primary & primary* *Podstawowe i mniej**	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.45
ASFRs for the total population *Współczynniki płodności dla całej populacji*	0.07	0.34	0.44	0.26	0.11	0.02	0.00	1.25

* Reference group – Less than primary and primary education.

* *Grupa referencyjna – Wykształcenie podstawowe i niepełne podstawowe.*

Source: Own calculations, Eurostat, CSO.

Źródło: Obliczenia własne, Eurostat, GUS.

showed that the age within which transitions to different model states were possible can be limited to the 12–34 years of age. The matrix of transition probabilities, for each age and sex group, was calculated using the structure of educational attainment in the year 2002, for the same age group in the period (t, $t+1$). Next, the education structure of the age groups in the period $t+1$ was slightly modified, on the basis of information from the LFS. Then, the Iterative Proportional Fitting procedure (IPF) was applied to calibrate the probabilities of possible flows[6]. The

---

The elimination of impossible transitions (for example from advanced educational level to less advanced educational level) allowed to reduce the number of estimated parameters and the data check. Additionally, no changes in educational attainment for the persons older than 35 years was assumed due to rare transitions observed.

[6] The algorithm of the Iterative Proportional Fitting (IPF) is taken from Norman (1999). In this publication Iterative Proportional Fitting is described as "mathematical scaling procedure which can be used to ensure that a two-dimensional table of data is adjusted so that its row and column totals agree with constraining row and column totals obtained from alternative sources. IPF acts as a weighting system whereby the original table values are gradually adjusted through repeated calculations to fit the row and column constraints" (Norman, 1999: 2).

starting values for this procedure were taken from the probability matrix obtained from the LFS panel data. The final step was the comparison of the obtained results with published aggregate statistics on graduates from different types of schools[7].

In the current version of the projection model, migration flows are not taken into account and the "zero net-migration scenario" has been introduced. Other versions of this assumption would be interesting to consider, but this is difficult because of:

1) the lack of reliable and detailed data on migration inflows and outflows by age, sex and educational attainment,
2) problems with definitions of migrants in the CSO data.

Therefore, adding migration assumptions would be extremely subjective and would require a more developed analysis of different scenarios. Such an analysis exceeds the aims and capacity of this paper.

PROJECTION RESULTS

This section presents the projection outcomes with the aim to demonstrate changes in the population structure by educational attainment. These results are also compared with the observed values from the years 2003–2006, to show how the projections perform in the short term. The reference to fertility and mortality assumptions of the Eurostat's official projection EUROPOP 2004 is also made, in order to assess, to what extent the future assumptions can be explained by current projections.

The departure point for an analysis of the consistency between projections and empirical observations, is the comparison of projection results for the period 2003–2006 with the official estimates published in the Statistical Yearbooks of Poland[8]. The official figures are based on the census data and on reports from educational institutions. The Labour Force Survey is another source of data on educational attainment, but as it was explained in the previous section, the results of this survey may be biased to some extent.

The comparison of the projection results for the years 2005 and 2006 with the actual data on the population structure by educational attainment shows that the projected values are quite similar to the values obtained from the Population Census 2002 and Labour Force Survey for the years 2004–2006 (Table 4).

The trends in population shares by educational attainment in the period 2002–2006 could be relatively well explained using the multi-state projection with constant assumptions, based on the 2002 data. This observation could suggest that the changes in mortality, fertility and probabilities of transition between education levels in this relatively short period of time were to small to generate major differences between the observed composition by educational attainment and the projection results.

---

[7] On the basis of the data from CSO (2003a, 2004a, 2004b).

[8] See CSO (2003b, 2004c, 2005, 2006, 2007a).

Table 4. The structure of the Polish population (aged over 13) by sex and educational attainment (in percents): Comparison of different data sources with the baseline projection

Tabela 4. Struktura ludności Polski (powyżej 13 lat) według wieku, płci oraz poziomu wykształcenia (procentach): porównanie danych z różnych źródeł z wynikami projekcji

Sex / Educational attainment *Płeć / Poziom wykształcenia*	2002			2004			2006		2010	2020	2030
	Official estimates* *Oficjalne szacunki**	LFS data*** *Dane BAEL****	Projection *Projekcja*	Official estimates** *Oficjalne szacunki **	LFS data*** *Dane BAEL****	Projection *Projekcja*	LFS data*** *Dane BAEL****	Projection *Projekcja*	Projection *Projekcja*	Projection *Projekcja*	Projection *Projekcja*
Males *Mężczyźni*											
Tertiary *Wyższe*	9.3	9.4	9.8	0.1	11.8	11.0	12.5	12.3	15.1	21.0	25.3
Post-secondary *Policealne*	1.6	1.4	1.6		1.4	1.7	1.6	1.8	2.0	2.5	2.7
Secondary vocational *Średnie zawodowe*	20.6	20.5	21.3	29.8	21.0	22.1	21.2	22.8	23.8	25.4	26.1
Secondary general *Średnie ogólne*	5.4	5.2	5.7		6.5	6.0	6.9	6.2	6.4	6.5	6.9
Basic vocational *Zasadnicze zawodowe*	30.1	35.0	30.7	30.0	33.9	30.3	31.8	30.0	29.5	27.7	24.5
Primary *Podstawowe*	28.0	27.2	27.9	26.6	24.1	26.0	25.2	24.3	21.1	15.5	12.8
Less than primary and no education *Mniej niż podstawowe i bez wykształcenia*	5.0	1.3	3.0	1.9	1.3	2.8	0.8	2.5	2.1	1.5	1.7

Sex / Educational attainment *Płeć / Poziom wykształcenia*	2002			2004			2006		2010	2020	2030
	Official estimates* *Oficjalne szacunki**	LFS data*** *Dane BAEL****	Projection *Projekcja*	Official estimates** *Oficjalne szacunki **	LFS data*** *Dane BAEL****	Projection *Projekcja*	LFS data*** *Dane BAEL****	Projection *Projekcja*	Projection *Projekcja*	Projection *Projekcja*	Projection *Projekcja*
Females *Kobiety*											
Tertiary *Wyższe*	10.4	10.8	11.1	0.1	13.1	13.0	15.0	14.9	18.8	27.0	33.0
Post-secondary *Policealne*	4.6	3.9	4.8	33.0	4.0	5.0	4.1	5.2	5.6	6.3	6.7
Secondary vocational *Średnie zawodowe*	18.7	19.0	19.4		19.4	19.7	19.4	20.0	20.4	21.1	20.9
Secondary general *Średnie ogólne*	11.7	11.8	11.9		12.9	12.0	12.5	11.8	11.4	10.2	9.5
Basic vocational *Zasadnicze zawodowe*	16.9	20.2	17.2	17.2	19.8	17.0	18.6	16.9	16.9	16.7	15.6
Primary *Podstawowe*	31.4	31.1	31.4	29.5	28.2	29.4	28.8	27.6	24.1	17.1	12.7
Less than primary and no education *Mniej niż podstawowe i bez wykształcenia*	6.3	3.2	4.2	3.1	2.7	3.9	1.7	3.5	2.8	1.6	1.6

* Estimates published in the Statistical Yearbook 2003, ** Estimates published in the Statistical Yearbook 2006, *** LFS (Labour Force Survey) data includes only the population 15 and over.

** *Szacunki opublikowane w Roczniku Statystycznym 2003,* ** *Szacunki opublikowane w Roczniku Statystycznym 2006;* *** *Dane BAEL (Badanie Aktywności Ekonomicznej Ludności) zawierają informacje o wykształceniu osób w wieku 15 lat i więcej.*

Source: CSO, LFS, own calculations.
Źrodło: GUS, dane BAEL, obliczenia własne.

The differences between the projected and observed values can be explained not only by the initial differences in the observed shares by education, but also by the changing preferences of pupils and students, which result in shifting education paths at the micro level, in comparison to the year 2002. These changes could account for the rapid increase of the share of people with tertiary education between the years 2002–2004, and a slightly faster than projected drop in the share of population with basic vocational education. Despite the fact that the projection slightly underestimates the pace of the increase in the share of people with tertiary education in the period 2002–2006, the values projected after 2006 show that in the following 25 years the share of people with tertiary education in the total population would be expected to double. The general trend of acquiring higher levels of education will lead to an increase in the shares of different types of secondary education, while the group of people with only basic vocational education will become smaller. According to the projection the share of people with only primary education or less education has dropped by about 4 percentage points between 2002 and 2006. This change in recent years was mainly triggered by the fading of the oldest cohorts with relatively high percentage of people with primary or less education acquired before the Second World War. The projection results show that over the next decades virtually all people from the new generations can be expected to attain a higher than primary level of education (Figures 1a and 1b).

Provided that the trends from the years 2002–2006 will be sustained, then over the next 25 years, Poland may increase the share of people with higher education in the population over 15 years from about 12% to about 25% in 2030. It is approximately the equivalent to that achieved in the United States[9] in the period 1980–2000.

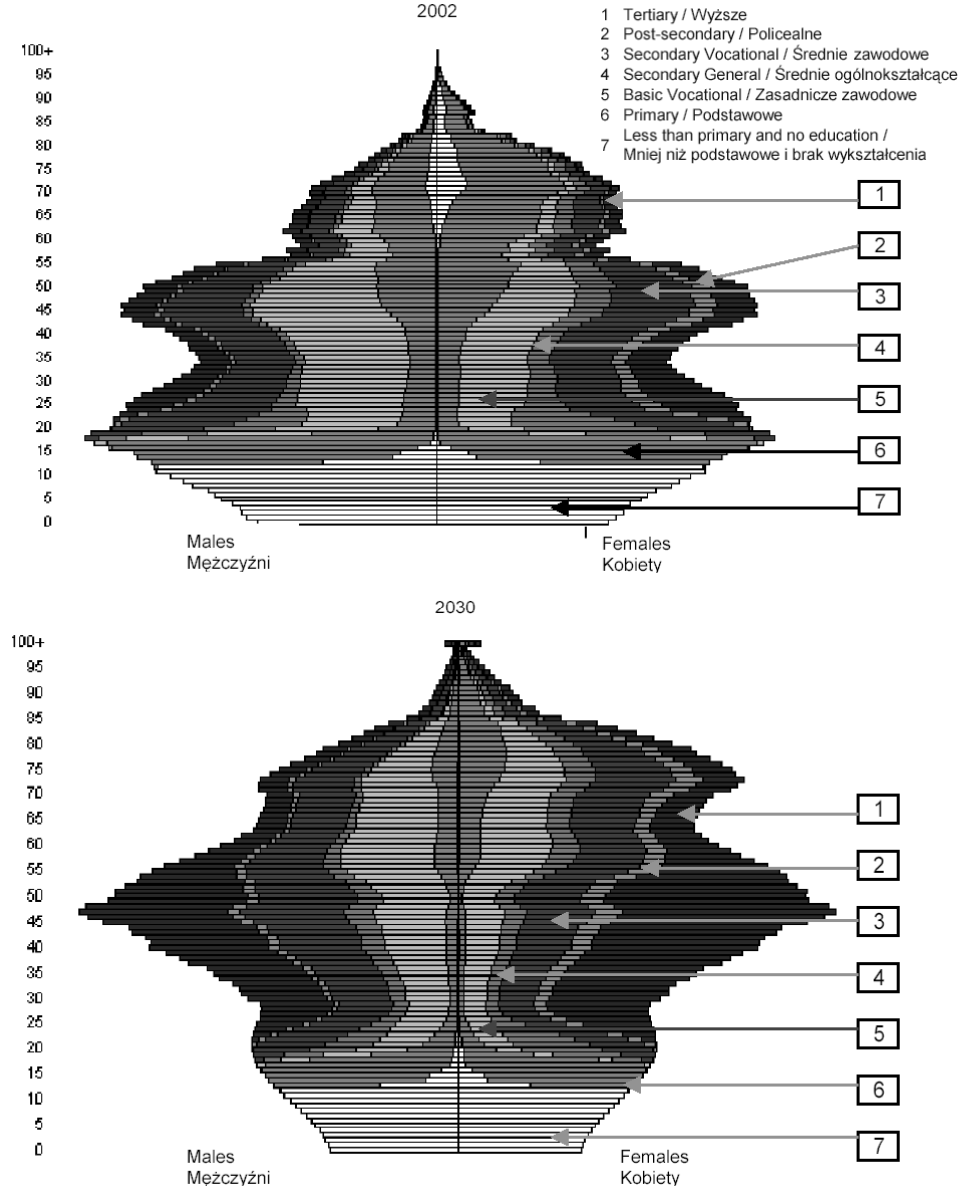
If the analysis was limited only to persons in the working-age group (15–59/64), the projection results would show an even stronger and more intensive improvement in the population structure by completed education. This is because the majority of the numerous and relatively less educated older generations (such as the "baby-boomers" born in the 1950s and 1960s), is going to reach retirement age (60 years for woman and 65 years for man) between 2010 and 2020. These generations will be replaced by numerous and relatively better educated persons born at the end of the 1970s and beginning of the 1980s, who are accomplishing their tertiary education in the current decade (2000–2010). Provided that the TFR remains close to the level observed in 2002, the generation of their children is expected to be far less numerous. At the end of the projection period, people with tertiary education will constitute over 35% of the working-age population. Projection results also indicate that the size of working-age population groups with different educational attainment will be strongly influenced by the shifts in the age structures (Table 5).

---

[9] According to the data from Barro and Lee (2000b) and their discussion in Barro and Lee (2000a).

Figures 1a, 1b. Demographic pyramids by educational attainment in 2002 (Population Census) and 2030 (projection)

Figura 1a, 1b. Piramidy struktury demograficznej ludności Polski z uwzględnieniem poziomów wykształcenia w roku 2002 (Spis Powszechny) i 2030 (projekcja)



Source: CSO and own calculations using LIPRO 4.0.

Źródło: GUS i obliczenia własne w programie LIPRO 4.0.

Table 5. Projected working-age population by educational attainment
Tabela 5. Wyniki projekcji zmian struktury ludności w wieku produkcyjnym według wykształcenia

Educational attainment *Poziom wykształcenia*	Projection (15–59/64) *Projekcja (15–59/64)*			
	2002	2010	2020	2030
Tertiary *Wyższe*	11.7%	19.3%	28.6%	35.5%
Post-secondary *Policealne*	3.8%	4.2%	4.4%	4.5%
Secondary vocational *Średnie zawodowe*	23.0%	24.0%	23.7%	22.6%
Secondary general *Średnie ogólne*	9.6%	9.3%	8.5%	8.4%
Basic vocational *Zasadnicze zawodowe*	27.8%	25.9%	22.7%	18.1%
Primary *Podstawowe*	22.9%	16.2%	11.3%	10.1%
Less than primary and no education *Mniej niż podstawowe i bez wykształcenia*	1.1%	1.0%	0.8%	0.8%
Working-age population (in thous.) *Ludność w wieku produkcyjnym* *(w tys.)*	25 662	26 140	23 798	21 932

Source: CSO, own calculations.
Źródło: GUS, obliczenia własne.

According to the projection, the size of the working-age population is expected to grow from about 25.7 million in 2002 to about 26.1 million in 2010. This growth will be a result of attaining maturity by the "echo" of the baby boom generation of the late 1970s and early 1980s. This increase would be then followed by a continuous decline, to about 23.8 million in 2020 and 21.9 million in 2030, caused by entering the working age by much smaller generations born in the period of fertility declines to very low levels (since the 1990s).

This initial upward and then declining trend will not be present in the group of people with only vocational education or less. The share of this group is projected to shrink during the entire projection period. On the other hand, the number of people with tertiary education is projected to increase throughout the projection horizon.

INFLUENCE OF THE CHANGES IN THE EDUCATION STRUCTURE ON SOME DEMOGRAPHIC INDICATORS

The possible influence of predicted changes in the population composition by educational attainment on future aggregate measures of fertility and mortality, life expectancy at birth (e_0) and total fertility rate (*TFR*), is analysed by comparing

obtained projections results with the results without taking into account the population heterogeneity by educational attainment. Here, the population projection prepared by the Eurostat will be referred to (EUROPOP 2004). Furthermore, the previewed rapid changes in the structure of population by education level are discussed in the context of possible influence on aggregate measures of fertility and mortality rates in the future. The first simulation regards the differences between population groups in mortality rates. The upward trend in life expectancy was observed in Poland over the past years, like in most European countries. According to empirical data, life expectancy in Poland in the period 2002–2006 increased by 0.5 years for men and 0.6 years for women (Table 6).

Table 6. Life expectancy at birth (e_0): comparison of observed values, assumptions of the Eurostat projection EUROPOP 2004 and the current projection

Tabela 6. Oczekiwana długość trwania życia w momencie narodzin (e_0): porównanie wartości zaobserwowanych, założeń projekcji Eurostatu EUROPOP 2004 oraz obecnej projekcji

Sex / data source *Płeć / Źródło danych*	Life expectancy at birth (e_0)						
	2002	2006	Difference *Różnica* 2006–2002	2010	2020	2030	Difference *Różnica* 2030–2002
Males *Mężczyźni*							
Observed (for 2002 and 2006) *Obserwacje (dla lat 2002 i 2006)*	70.4	70.9	0.5				
EUROPOP 2004 baseline projection *Bazowa projekcja EUROPOP 2004*	70.4	71.0	0.6	72.0	74.5	76.8	6.4
Current projection by education *Obecna projekcja wg wykształcenia*	70.4	70.8	0.4	71.1	71.7	72.6	2.2
Females *Kobiety*							
Observed (for 2002 and 2006) *Obserwacje (dla lat 2002 i 2006)*	78.8	79.6	0.8				
EUROPOP 2004 baseline projection *Bazowa projekcja EUROPOP 2004*	78.8	78.9	0.1	79.6	81.3	82.8	4.0
Current projection by education *Obecna projekcja wg wykształcenia*	78.8	79.7	0.2	79.1	79.9	80.8	2.0

Source: CSO, Eurostat, own calculations.
Źródło: GUS, Eurostat, obliczenia własne.

Comparing these figures with the assumptions of the baseline scenario of the EUROPOP 2004, it appears that the decline of mortality rates for men were overestimated and for women were underestimated. In this context it should be mentioned, that the population heterogeneity by education is explicitly not taken into account in the Eurostat's population projection.

Had the heterogeneity of the population by education level been taken into account, the changes in population structure by education level might have explained about 0.4 years of life expectancy for men and about 0.2 years for women, even if mortality rates within the education groups would stay constant at the level from the year 2002. It should also be mentioned, that according to the assumptions of the Eurostat projection, the increase in female life expectancy between 2002 and 2006 amounted to about 0.1 year, which was lower than what was reflected by the changes in the structure of population by educational attainment.

The projected changes of population by educational attainment can be used to calculate how the changes in the population distribution by education level can influence the life expectancy of the total population. Educational attainment seems to reflect lifestyle, socio-economic status, health and other mortality-relevant factors. The increase in the share of people with higher levels of education should lead to higher life expectancy of the whole population. As it could be seen from Table 2, the highest life expectancy at birth was noted for persons of tertiary education attained after turning 24 years of age. For them, the e_0 amounted to about 79.8 years for males and 86.1 years for females, which is the maximum level of life expectancy attained by all population groups under study.

In other words, if all people aged 24 years and more attained tertiary education, higher levels of life expectancy could be reached. However, the assumptions of our projected changes in the population structure are not so strong. On the other hand, including the education differences in the analysis of mortality resulted in the increase of male life expectancy up to 72.6 years until the year 2030 (thus, by about 2.2 years). This increase was only due to the projected improvement in educational attainment. The same changes in the education structure of women could contribute to an increase of life expectancy to about 80.8 years (thus, by about 2 years). By comparing these results with the assumptions of the EUROPOP 2004, it appears that the changes in the population education composition may explain about half of the increase in e_0 for females, and about 30% for males, as projected by the Eurostat.

According to the data from the year 2002, presented in table 3, the attainment of higher than basic vocational education decreases age-specific fertility rates in the age groups of 15–29 and 35–49 years. Only fertility rates of women with tertiary education aged 30–34 exceed those of other groups. This concentration of fertility could be the result of both fertility and labour market decisions that were made by the previous cohorts of women with tertiary education after having accomplished the education process. These assumptions strongly influence the simulation results of future changes in TFRs (Table 7).

Table 7. Total Fertility Rate (TFR): comparison of the observed values, assumptions of the Eurostat projection EUROPOP 2004 and the current projection

Tabela 7. Ogólny współczynnik dzietności: porównanie wartości obserwowanych, założeń projekcji Eurostatu EUROPOP 2004 i obecnej projekcji

Source of data *Źródło danych*	Total Fertility Rate (TFR)						
	2002	2006	Difference *Różnica* 2006–2002	2010	2020	2030	Difference *Różnica* 2030-2002
Observed values (for 2002 and 2006) *Obserwacje (dla lat 2002 i 2006)*	1.25	1.27	0.02				
EUROPOP 2004 baseline projection *Projekcja EUROPOP 2004*	1.25	1.18	-0.07	1.19	1.42	1.58	0.33
Current projections by education *Obecna projekcja wg wykształcenia*	1.25	1.20	-0.05	1.21	1.21	1.21	-0.04

Source: CSO, Eurostat, own calculations.
Źródło: GUS, Eurostat, obliczenia własne.

The simulation shows, that in the period 2002–2006 the TFR should have dropped from 1.25 to 1.20. The projected value was very similar to that assumed in the EUROPOP 2004 projections, however in 2006 the observed TFR was not as low as projected. This phenomenon could be explained by the effect of first birth postponement, which seems not to be taken into account both, both in the simulations and in the Eurostat projections.

As we can see for the period after 2006, the TFR assumptions in the EUROPOP 2004 and the simulation results of the impact of changes in the education composition on fertility are completely different. In the simulation, the changes over the successive periods are smaller, with a slight increase occurring after 2010, which reflects the fact that numerous generation of highly-educated women will reach the age 30 years and over.

The inaccuracy of the simulation result for 2005 requires some explanation. The analyses presented in this chapter are based on cross-sectional indicators. By excluding changes in the educational structure it is assumed that subsequent cohorts are similar to their predecessors. However, the analyses of changes in fertility showed that this assumption was too strong. Longer participation of cohorts in the education system could contribute not only to a decrease in fertility rates, but also to the postponement of planned births. This could result in an increase of fertility rates in older age groups, as it was observed in 2005. Recuperation at older ages contributed to a slow fertility increase observed in Poland since 2004.

CONCLUSIONS

There are many reasons, as to why knowledge of the present and future education compositions of the population is important. Multi-state models create an opportunity for including education in population projections and have many advantages over other methods usually used to project the educational structure. The multi-state approach makes it possible, among others, to analyse the transitions between various educational attainment levels, and to account for the consequences of the current population structure by education on the future structure. The projection results could not only provide information on the future changes of the education composition itself, but also allow to include important source of the population heterogeneity in the projections, and explain to some extent the possible future changes of aggregate demographic indicators.

According to the author's knowledge, this study provides the first projections by educational attainment of the population of Poland. The results show that if the probabilities of transitions between educational levels were stable and equal to those from the period 2002–2003, the share of persons with more advanced educational attainment would increase remarkably. The share of people with tertiary education in the working-age population would grow from about 12% in 2002 to 29% in 2020 and over 35% in 2030, what could have significant consequences at least for the labour market. Comparisons of the projection results with the observed changes in the population structure by educational attainment for the period 2003–2006 demonstrated that the projections under constant transition rates for education produced quite similar values, and thus in the short term could possibly be useful also as forecasts.

The projection results were also applied to assess to what extent the changes of the population composition by education could influence the values of demographic indicators. Providing that the differences in mortality rates between different education groups remain constant in the future, the life expectancy of the whole population may increase by 2.2 years for men and 2.0 years for women in the years 2003–2030 only due to the changes in the education structure. However, this increase would be still insufficient to catch up even with recent values of life expectancy prevailing in the majority of Western European countries.

The attempt to explain future developments in the TFR appeared to be less successful. Here, the turning point of 2004 in the fertility changes seems to be one of the reasons beside the fact that in this case a cohort analysis would be more appropriate. The cross-sectional fertility rates used in the projection do not reflect the distinct changes in the timing of fertility between generations of women. On the other hand, according to the projection, the negative influence of higher education boom on the *TFR* was an important factor over the last 5 years, but it might not cause any further decrease in fertility rates in the future.

The projection model presented requires some further improvements, such as including migration, redefinition of the initial population or possible update of the parameters to the observed changes in the transition probabilities. However,

even with its simplistic assumptions the model seems to deliver a quite reliable and detailed output, which could be useful for various demographic and economic analyses.

REFERENCES

Barro R.J., Lee J., 2000a, *International data on educational attainment; updates and implications*. CID Working Paper No. 42, Center for International Development, Harvard University, Cambridge, MA.

Barro R.J., Lee J., 2000b, International Data on Educational Attainment – Human Capital Updated Files, http://www.cid.harvard.edu/ciddata/ciddata.html (access date: 28.11.2008).

Becker G., Murphy K., Tomura R., 1990, *Human capital, Fertility and Economic Growth*. "The Journal of Political Economy", Vol. 98, No. 5, 12–37

Bolesławski L., 2004, *Demographic projections of Poland 2003–2030.* Central Statistical Office, Warsaw.

CSO, 2003a, *Higher education institutions and their finances 2002.* Central Statistical Office, Warsaw.

CSO, 2003b, *Statistical yearbook of Poland 2002*. Central Statistical Office, Warsaw.

CSO, 2003c, *Life expectancy in Poland in the year 2002*. Central Statistical Office, Warsaw.

CSO, 2003d, Ludność. Stan i struktura demograficzno-społeczna 2002, Central Statistical Office, Warsaw

CSO, 2004a, *Education in the school year 2003/2004*. Central Statistical Office, Warsaw.

CSO, 2004b, *Higher education institutions and their finances 2003.* Central Statistical Office, Warsaw.

CSO, 2004c, *Statistical yearbook of Poland 2003*. Central Statistical Office, Warsaw.

CSO, 2005, *Statistical yearbook of Poland 2004*. Central Statistical Office, Warsaw.

CSO, 2006, *Statistical yearbook of Poland 2005*. Central Statistical Office, Warsaw.

CSO, 2007a, *Statistical yearbook of Poland 2006*. Central Statistical Office, Warsaw.

CSO, 2007b, *Baza danych demografia, Tablice Statystyczne z Demografii za lata 1999–2006*, http://www.stat.gov.pl/demografia/index.html (date of access: 30.10.2007).

Cohen D., Soto M., 2001, *Growth and Human Capital: Good Data, Good Results*. CEPR Discussion Paper No. 3025, Center for Economic Policy Research, London.

Florczak W., 2008, *Kapitał ludzki w empirycznych modelach wzrostu*. "Ekonomista", No. 2/2008, 169–200.

Goujon A., Wils A., 1996, *The importance of Education in the Future Population. Global Trends and Case Studies on Cape Verde, Sudan and Tunisia*. IIASA Working Paper WP-96-138, International Institute for Applied Systems Analysis, Laxenburg.

Huisman M., Kunst E., 2004, *Socioeconomic inequalities in mortality among elderly people in 11 European populations*. "Journal of Epidemiology and Community Health", Vol. 58, 468–475.

Jóźwiak J., Kotowska I.E., Kowalska A., 2000, *Procesy demograficzne, rynek pracy, edukacja*, [in:] Ekonomiczne i społeczne efekty edukacji, Instytut Problemów Współczesnej Cywilizacji, Warszawa.

Kotowska I.E., 1994, *Population Dynamics in Poland, 1950–2050: Internal Migration and marital Status Changes*. IIASA Working Paper WP-94-74, International Institute for Applied Systems Analysis, Laxenburg.

Kotowska I. E., Sztanderska U., Wóycicka I., 2007, *Aktywność zawodowa i edukacyjna a obowiązki rodzinne w Polsce w świetle badań empirycznych*. Wyd. Scholar, Warszawa.

Kupiszewska D., Kupiszewski M., 2005, *A revision of traditional multiregional model to better capture international migration: The MULTIPOLES model and its application*. CEFMR Working paper 10/2005, Central European Forum for Migration and Population Research, Warsaw.

Kupiszewski M., Bijak K., Saczuk K., Serek R., 2003, *Komentarz do założeń prognozy ludności na lata 2003–2030 przygotowanej przez GUS*. CEFMR Working Paper 3/2003, Central European Forum for Migration and Population Research, Warsaw.

Lutz W., Goujon A., Doblehammer-Reiter G., 1998, *Demographic dimension in forecasting*, [in:] Lutz W., Vaupel J., Ahlburg D. (eds.), *Frontiers of Population Forecasting*, "Population and Development Review", Supplement to Vol. 24, 42–58.

Lutz W. Goujon A., 2001, The *World's Changing Human Capital Stock: Multi-State population Projections by Educational attainment*. "Population and Development Review", Vol. 27, No. 2, 323–339.
Mankiw N. G., Romer D., Weil D. N., 1992, *A Contribution to the Empirics of economic Growth*. "The Quarterly Journal of Economics", Vol. 107, No. 2, 407–437.
Matysiak A., Nowok B., 2007, *Stochastic forecast of population of Poland, 2005–2050*, "Demographic Research" Vol. 17, No. 11, 301–338.
MEN, 2008, *Co to jest klasyfikacja ISCED*, http://www1.men.gov.pl/content/view/10446/ (date of access: 27.11.2008).
Norman P., 1999, *Putting Iterative Proportional Fitting on the research Desk*. University of Leeds Working Paper 99/03, University of Leeds, Leeds.
Paradysz J., 1981, *Wielostanowa analiza demograficzna*, „Studia Demograficzne", No. 1/63, 75–89.
Paradysz J., 1988, *Demografia wielowymiarowa w Polsce*, „Studia Demograficzne", No. 2/88, 131–141.
Preston S., Heuveline P., Guillot M., 2001, *Demography. Measuring and Modelling Population Processes*. Blackwell Publishers Ltd., Oxford.
Rogers A., 1975, *Introduction to Multiregional Mathematical Demography*. John Willey, New York.
Rogers A. (ed.), 1980, *Essays in multistate demography*. "Environment and Planning A", Vol. 12, No. 5, 485–622.
Sakson B., 2002, *Wpływ „niewidzialnych" migracji zagranicznych lat osiemdziesiątych na struktury demograficzne Polski*, „Monografie i Opracowania", nr 481, SGH, Warszawa
Shkolnikov V., Andreev E., Jasilionis D., Leinsalu M., Antonova O., McKee M., 2006, *The changing relation between education and life expectancy in central and eastern Europe in the 1990s*. "Journal of Epidemiology and Community Health", No. 60, 875–881.
Shryock H.S., Siegel J.S., 1976, *The Methods and Materials of Demography*, Academic Press, New York.
UN, 1958, *Multilingual Demographic Dictionary. English Section*, United Nations Population Studies, No. 29.
Van Imhoff E., Keilman N., 1991, *LIPRO 2.0: An Application of a Dynamic Demographic Projection Model to Household Structure in The Netherlands*, Netherlands Interdisciplinary Demographic Institute, The Hague.
Van Imhoff E., 1994, *LIPRO USER'S GUIDE Version 3.0*. Netherlands Interdisciplinary Demographic Institute, The Hague.
UNESCO, 2006, *International Standard Classification of Education ISCED 1997*, http://www.uis.unesco.org/TEMPLATE/pdf/isced/ISCED_A.pdf (date of access: 27.11.2008).
Willekens F.J., Drewe P., 1984, *A multiregional model for regional demographic projection*, [in:] H. ter Heide, F.J. Willekens (eds.), *Demographic Research and Spatial Policy. The Dutch Experience*, Studies in Population, Academic Press, London.
Wróblewska W., 2006, *Analiza umieralności w Polsce w latach 1970–2003. Dekompozycja zmian w oczekiwanym trwaniu życia noworodków*, „Studia Demograficzne", No. 1/149, 28–48.

THE MULTI-STATE PROJECTION OF POLAND'S POPULATION BY EDUCATIONAL ATTAINMENT FOR THE YEARS 2003–2030

The first population projection by education level of the Poland's population until the year 2030 is presented. The projecton is based on the multi-state projection model LIPRO developed by E. van Imhoff and N. Keilman (1991) and the LIPRO 4.0 software.

The initial population as well as the model parameters were calculated for the year 2002 on the basis of the National Population Census data and the Labour Force Survey data. The projection was prepared under an assumption on the constant parameters up to the year 2030.

The differences in mortality and fertility by education attainment were also taken into consideration.

The projection results for the years 2002–2006 were compared with the observed values to check the assumptions formulated.

The projection results show that the increasing enrolment at the tertiary and secondary levels of education among the baby boomers born in the mid-1970s and the early 1980s will dramatically change the population composition by education in the next two decades. The percentage of people with tertiary education in the working age population (15-59/64) will increase from 12% in the year 2002 to 35% in the year 2030. Moreover, the analysis was performed to demonstrate changes in the life expectancy at birth (e_0) and the total fertility rate (TFR) which can be attributed to the changing population composition by education and existing differences in mortality and fertility by education, to be kept in the future. Until 2030 the life expectancy increases by about 2.2 years for men and 2.0 years for women only because of the shifts in the education composition. On the other hand, this factor was found as not contributing to the fertility changes in the next years.

Key words: population projection, education attainment, fertility and mortality differences by education