UNIVERSITY OF ECONOMICS,

PRAGUE

MASTER´S THESIS

2019 **Polina Bezushchenko**

University of Economics, Prague

Faculty of Informatics and Statistics

Study programme: Quantitative Methods in Economics

Field of study: Official Statistics

ACCURACY EVALUATION OF POPULATION PROJECTIONS IN THE CZECH REPUBLIC

Master´s thesis

Author: Bc. Polina Bezushchenko

Supervisor: Ing. Petr Mazouch, Ph.D.

Prague, April 2019

DECLARATION:

I hereby declare that I am the sole author of the thesis entitled "Accuracy evaluation of population projections in the Czech Republic". I duly marked out all quotations. The used literature and sources are stated in the attached list of references.

In Prague on ....

Signature

Polina Bezushchenko

ACKNOWLEDGEMENT

I hereby wish to express my appreciation and gratitude to the supervisor of my thesis,
Ing. Petr Mazouch for his help and support;
to the supervisor of my project during the internship,
Mgr. Terezie Štyglerová for her direction and insights;
to *Mgr. Michaela Němečková* for providing with the data;
to *RNDr. Luděk Šídlo* for explaining the methodology;
to *RNDr. Boris Burcin* for providing with the data and explaining the methodology.

Title: Accuracy evaluation of population projections in the Czech Republic

Author: Polina Bezushchenko

Study program: Official Statistics

Supervisor: Ing. Petr Mazouch, Ph.D.

Abstract:

Nowadays, population projections are widely used at the different levels of national planning as well as by businesses. For the last decade a lot of new projections have been released for the population of the Czech Republic up to 2101. The accuracy evaluation of the current projections can help policymakers to understand how the future population may unfold. Also, knowing the errors of the projections, the future projections can be improved. In this thesis several current population projections are evaluated against the reality with the help of the Keyfitz's "Quality of Prediction Index" and the Mean Absolute Percentage Error. The evaluation was conducted for the projections published by the Czech Statistical Office, Eurostat, the United Nations and by individual researches Boris Burcin and Tomáš Kučera. The basic results and important findings are presented together with the description of the individual projections. The results reveal that the most accurate age groups are 10-19 and 60-69; the least accurate age groups besides old ages are 0-9 and 20-39. The most problematic parameters are net migration and life expectancy at 65. The accuracy of the prediction seems to be very high during the first 2 years after the publication not exceeding the deviation of 1%. The error starts to rise after 4 years elapsed from the projections' release exceeding the deviation of 1% and more. The projection of Eurostat seems to be the most accurate one. To the contrary, the least accurate projection belongs to Burcin and Kučera.

Keywords: population projection, population forecast, forecast accuracy, accuracy evaluation, demographic development, Czech Republic, Keyfitz's "Quality of prediction index", Mean Absolute Percentage Error

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ABS	Australian Bureau of Statistics
APA	American Planning Association
ARIMA	Autoregressive Integrated Moving Average model
B&K	projection of Boris Burcin and Tomáš Kučera
CSU	projection of Czech Statistical Office
CZSO	Czech Statistical Office
ESRI	Environmental Systems Research Institute
IIASA	International Institute for Applied Systems Analysis
MALPE	Mean Algebraic Percent Error
MAPE	Mean Absolute Percent Error
MAPE-R	Mean Algebraic Percent Error Re-scaled
MEDAPE	Median Absolute Percentage Error
MPE	Mean Percentage Error
MSPE	Mean Square Percentage Error
NRC	National Research Council of Washington
PRB	Population Reference Bureau
RMSPE	Root Mean Square Percentage Error
UN	United Nations
WB	World Bank
WPP	World Population Prospects, projection of United Nations

1 INTRODUCTION

Not long-ago population projections became a significant part of demographic statistics. First of all, demographers started to be concerned about the rapid increase in the world population during the second demographic transition. People started to worry there may be too many people in the future on the earth which may exceed the earth's ability to feed, clothe, and house the human beings (Zhao, 2010). At the national level, people were concerned about the economic, social, political and environmental consequences of population growth and demographic change (Zhao, 2010). That is how constructing population projections became an essential activity for the demographers all over the world. "*The earliest systematic global population projection dates to Notestein in 1945, although many national level projection efforts began over half a century earlier*" (O'Neill, Balk, Brickman & Ezra, 2001).

Since the fluctuating population size of a country may influence all areas of environment a lot, there is no doubt that monitoring and evaluating population projections' accuracy is a crucial thing. For instance, "*trends in population by age are needed to projection the demand for education, and to plan the provision of education at all levels*" (Billari, Graziani, Melilli, 2011). Or, "*population projections of a century or more are frequently used by climate change researchers for the estimation of future risk and the analysis of policy options*". (Smith, 2011) Similarly, the proportion of elder people or old-aged dependency ratio required for planning the state budget for the pension system and regulating the retirement age.

Many people and users of population projections believe that projections are precisely accurate. There are less people who are aware of the errors that are inherent to projections, especially for the local and small area. However, "*large errors, such as 10% or more after just 10 years into the projection, are common for local and small area populations, as shown by earlier research on projection error and accuracy like Isserman in 1977, Rayer in 2008, Smith and Shahidullah in 1995, Tayman in 1998, Wilson and Rowe in 2011*" (Wilson, Brokensha, Rowe, Simpson, 2017).

In the Czech Republic there are several authorities that are engaged into the development of population projections. First of all, it is a national statistical agency of the Czech Republic – Czech Statistical Office that is constantly elaborating its own projections. There are some academic projections that are prepared by local demographers, professors of the Charles University, Boris Burcin and Tomáš Kučera that are working out on the population projections for the private usage. Also, the future population size of the Czech Republic is projected by several international statistical agencies, like Eurostat and the United Nations. In addition, there are some special predictions done by the University of Economic in Prague, called prognosis of Human Capital, that estimate not the population structure itself but the educational structure of population.

The accuracy evaluation of the past population projections was conducted just once. However, the importance of the population projectioning and its evaluation is increasing nowadays in the Czech Republic mainly due to the rapidly ageing of the population year by year. The problem is that it is crucial to be aware of how fast the population will age far in the future, especially for the reform of the pension system.

Following this introduction, chapters “Motivation and aim”, “Theoretical background”, “Production of population projections”, “Methodology of accuracy evaluation” will be presented in the theoretical part of the work finishing with short description of the Czech Republic demography together with the overview of the current and past projections. In the practical part, section “Data” describes the data, the data collection, and how the data was processed. The chapter “Results” presents the main trends of the projections in the Czech Republic; the characteristics of the individual projections; the description of the parameters and the comparison of age groups; and the overall results of the analysis with the deviation intervals and errors. The main results and findings together with recommendations and improvements are recalled in the chapter “Discussion and Conclusion”.

2 MOTIVATION AND AIM

The author of this work has passed the mandatory internship at the Czech Statistical Office as an intern of the department of Population Statistics under the supervision of Terezie Štyglerová. The project assigned during the internship was related to the accuracy evaluation of the population projections in the Czech Republic and comparison of the accuracy results of the individual projections. This project provided the author with an insight to expand the topic and develop the project into the Master's thesis.

More than that, such kind of work is needed for the demography of the Czech Republic. There are several authorities that are constantly developing population projections for the population of the Czech Republic; however, the accuracy evaluation of the projections was done historically just once. This evaluation was done by Klára Tesárkova and Luděk Šídlo within the framework of the demographic journal "Demografie". This is the only one officially published work in 2009, where Šídlo and Tesárkova discuss the current population projections, compare them among each other and judge their accuracy and quality. They used such quantitative methods in his work as the Keyfitz's "Quality of prediction index", the Theil`s index U and the evaluation method based on the principle of APC models. The results of this work will be discussed in the chapter 7.

Also, there is one thesis written by the student of the Charles University in Prague Lenka Šmejkalová and supervised by Tomáš Kučera in 2011, which was dedicated to the accuracy evaluation of the population projections of Plzeň town. In her work she used the quantitative methods of the Keyfitz's "Quality of prediction index", the Mean Percentage Error and the Root Mean Squared Percentage Error.

Nowadays, we can find 9 new projections that have been issued (but not all of them were publicly published) for the last decade and that have not been evaluated at all. However, the current situation of the demography of the Czech Republic requires to understand how the future population may unfold. Specifically, it is needed due to the rapid population ageing which is followed by the reform of the pension system. Thus, the main goal of the thesis is the post-evaluation of accuracy of the population projections in the Czech Republic that were published in 2009 and later and comparison of the individual projections between each other for the period from 2009 to 2017. The goal is intended to be reached with the help of 2 different methods of the accuracy evaluation: the Keyfitz's "Quality of prediction index" and the Mean Absolute Percentage Error.

The outcomes of the study are supposed to illustrate the most common errors and deviations of current projections to avoid them in new sets of assumptions for the future projectioning. In this work, the author would like to evaluate the accuracy of several projections and find out the most accurate and the least accurate projection. Also, to detect what are the most problematic components of the projections, worst age groups and worst parameters, and what stands behind it; and, finally, reveal the trends of population projections produced for the population of the Czech Republic.

3 THEORETICAL BACKGROUND

Broadly speaking, population projection is an estimate of the future development of the numerical state and structure by sex and age of a particular population. They "*deals with computations of future projection size and characteristics based on assumptions about future trends in fertility, mortality and migration*" (Planning Tank, 2017).

Population projections are belonged to the basic group of predictions of the population and demographic statistics. According Burcin & Kučera (2010), from time perspective, population development is a long-term process because both the numerical state and the composition by age and sex of each human population reflect the decades and sometimes centuries of development. They also stress out that population projections form the basis and the result of the demographic reproduction processes of a population, such that it represents the main link between the past and the future effects of reproductive forces and, at the same time, are the symbol of the continuity of population development. Interestingly, it was mentioned by National Research Council of Washington [NRC] (2000) that "*population projections are the demographic outputs most used by non-demographers and most neglected by population scientists*".

It has become possible to predict populations because current population trends, changes and structure can be monitored, and this knowledge can be used for projecting population for future periods. Thus, the first task of making projections is "*assessing the plausibility of current demographic estimates and choosing appropriate assumptions about future trends*" (NRC, 2000). The second task is to monitor population changes, which are in fact primarily based on changes in births and deaths that are influenced by social and economic factors (Renkou, 1980). All this together gives a good background for demographers to be able to predict "*a certain period's total population, age and sex structure, the number of births and deaths, and migration*" (Renkou, 1980).

Population projections can have different characteristics as, for example, length of time horizon, output variables, final usage and the coverage area. They can be produced for local areas, like counties and cities, or for the entire population of the planet. O'Neill et al. stated in their "Guide to Global Population Projections" that shorter time horizons are typical for local-area projections, usually less than 10 years, while national and world projections can describe the population several decades into the future. Also, they compared short and long-term projections by output variables and concluded that "*long-term projections typically produce a more limited number of output variables, primarily population broken down by age and sex; and, in contrast, projections for smaller regions often include other characteristics as well, which might include educational and labor force composition, urban residence, or household type*" (O'Neill et al., 2001).

3.1 Projection vs. Forecast

There are two terms that are commonly used in various studies. It is necessary to stress out the distinction of those terms as they serve significantly different wording and output in demography. In several studies on population projections, the word 'projection' is used interchangeably with the word 'forecast'. Forecasts are "*often defined as predictions; projections are simply the outcome of calculations based on specified input data assumptions, and may be likely, implausible, or illustrative of extreme scenarios*" (Wilson et al., 2017).

Probably, people's confidence in accuracy of the projections may come from the ignorance of the difference between the terms 'projection' and 'forecast'. Many users of population projections, who treat them as forecasts, believe that projections lose their credibility raising more and more uncertainty. However, projections are never actual forecasts. "*Population forecasts are often referred to as 'projections', and this is absolutely understandable, as they are projecting something into the future*" (Capuano, 2015). But the term 'projection' implicates that "*it is a continuation of current trends and 'projection' implies something more sophisticated*" (Capuano, 2015).

Table 1 Difference of projections and projections

	Projection	**Forecast**
Nature	Simply indicates a future value for the population if the set of underlying assumptions occur.	The assumptions represent expectations of actual future events.
Type of information	Indicates what future values for the population would be if the assumed patterns of change were to occur. It can vary the current levels of overseas migration, births and deaths to provide differing scenarios, but in the end still a prediction is based on a trend. It is not a prediction that the population will necessarily change in this manner.	Forecasts speculate future values for the population with a certain level of confidence, based on current and past values as an expectation (prediction) of what will happen.
Example	A population has grown at 3% p.a. for the past 10 years from 10.0 to 10.3 million people, so according to the trend it will continue to grow at this rate and will equal to 10.6 million in 10 years.	A population has grown at 3% p.a. for the past 10 years from 10.0 to 10.3 million people, so the current trend is taken into account. However, according to conditional expert opinion the population size is anticipated to decrease for the future 10 years due to environmental changes and unnecessity of having children in attitude of families of present generation.

Sources: Australian Bureau of Statistics, 2013; Capuano, 2015

The main principles of the two terms are extracted from the two sources: "*The population experts on population projections, forecasts and weather*" (Capuano, 2015) and Australian Bureau of Statistics [ABS]; and presented in the table 1. Both wordings include data analysis, but "*the key difference between forecast and projection is the nature of the assertion in relation to the* *assumptions occurring*" (ABS, 2013). Also, the table brings the simple examples for better understanding their natures.

Taking into account that the data presented in the analysis is not population forecasts but the actual projections of the Czech Republic's population. Hence, the term 'projection' is preferred and chosen for the work to avoid the interchangeability and misuse.

3.2 Usage of population projections

Population projections are widely used at any levels of planning, budgeting and analytics, for example for water and food use, or provision of public services like education and health. "*Most of the important decisions about major land uses and services are derived from population estimates: the demand for water, power and waste disposal facilities; housing, open spaces and schools; supply of labor; spending power available for the retail trade; enlarging a power plant; or revising local bus routes*" (Planning Tank, 2017). Different statistical authorities rely on population projections when executing their significant reforms and implementing policies. For instance, Czech Statistical office is currently using population projections of the Czech Republic to conduct the reform of the pension system in the country. For that, they need to know what the portion of people aged 65+ yeas will be in the future and how fast this portion will grow, such that they could get the idea of how much the retirement age should be postponed.

The information about the future demand and supply is important not just for government, but also for businesses that may use population projections to plan for the potential future expansion or reduction of the production based on the target population, or even opening a new business.

Some users prefer projections with various scenarios, and for other groups it is more important to utilize just one but most likely scenario. Also, the users differ in their preferences about the time horizon. On the one hand, "*the policy community, including advocacy groups, often would ask to a single most likely scenario, including projections that reflect the influence of policy*" (O'Neill et al., 2001). On the other hand, "*global change researchers often use projections as exogenous inputs to studies on topics such as energy consumption, food supply, and global warming. These studies usually require projections with long time horizons and range of scenarios rather than a single most likely projection*" (O'Neill et al., 2001).

Another question is how to improve the usage of population projections by program planners and policymakers. According to Population Reference Bureau [PRB] (2001) there are several steps that can be launched to make projections more reliable and useful:

- Understand the causes of uncertainty in population projections and the implications of this uncertainty for plans and policies that span different time horizons and target specific population groups;
- Contribute to national and international efforts to collect more accurate demographic data — which would lead to more accurate assumptions about fertility, mortality, and migration and better projections;
- Cooperate with national and international research efforts to develop more accurate projections by supporting organizations that investigate better projection methodologies, the demographic effect of HIV/AIDS, the effect of policies and programs on fertility trends, and similar topics.

4 PRODUCTION OF POPULATION PROJECTIONS

4.1 Authorities that produce population projections

"*Although many national level projection efforts began over half a century earlier, the earliest systematic global population projection dates to American demographer Frank W. Notestein in 1945*" (O'Neill et al., 2001). He contributed to the development of demography significantly by outlining economic and social factors that influence the population growth. The United Nations had become a leader of population projectioning and dissimilation of its results since the 1950s. "*Later efforts have been undertaken by three other institutions: The United States Census Bureau, the World Bank, and the International Institute for Applied Systems Analysis*" (O'Neill et al., 2001). Nowadays, mostly national statistical agencies or governments are responsible for the production of population projections for their own countries and regions. Beside it, there are several international organizations that project populations for the world and individual countries. Also, there is a place to be for individual researches that are likely not intended to undertake "*global long-run population projections. However, individual researchers have tended to create projections at the national-level (or below) and at this level have made significant contributions to varying methodologies*" (O'Neill et al., 2001).

Among best-known and widely used international organizations with their own statistical database we can find the United Nations [UN], the World Bank [WB], International Institute for Applied Systems Analysis [IIASA], and Eurostat. The UN issue global and national projections and revise them on a regular basis, and, in fact, they are the most widely used projections all over the world. "*Many national governments, international agencies, the media, researchers, and academic institutions rely on UN projections. The WB, the IIASA and Eurostat also prepare population projections for the world, major regions, and, especially the WB and Eurostat, for individual countries. World Bank projections generally are used for planning and for managing projects, while IIASA projections have been used primarily to assess various projection assumptions and methods*" (PRB, 2001). Eurostat is the major database for the European countries, and it produces projections at national level comprises data for all EU-28 Member States including data for Norway. The methodology used by each of those international authorities slightly differs. This involves setting "*varying assumptions about the future demographic trends and starting with slightly different estimates of current population size*" (PRB, 2001).

4.2 Parameters and Approaches

The future population trend is determined by the interplay of mortality, fertility and migration. This is the main set of assumptions that must be cautiously and properly chosen, since it not just jointly draws the resulting projection but also claims the future reliability if it. All these variables contribute to the population growth with which help future population is calculated. "*Demographers base fertility and mortality rates on birth and death statistics*" (Dotson, 2018). Further, mortality and fertility split into age-specific rates and mortality rates define life expectancy at different ages. As was said by Caswell & Gassen (2015), "*Such scenario-building is a kind of perturbation analysis, quantifying the effects of large changes imposed on many vital rates simultaneously, but the number of possible scenarios is effectively infinite*".

As was mentioned in the previous chapter, the assumptions used for projecting a population involve that the recent demographic trends will continue, they are not the exact predictions. Thus, there are some challenges that population projections may face. The most straightforward challenge is that "*recent-trend projections that do not tend to account for other events that could change the shape of population growth. For example, such scenarios as conflict, an epidemiological disaster, natural disasters and extreme weather events, and food scarcity are more pressing in the context of climate change*" (Dotson, 2018). Those potential factors aggravate the reliability of population projections especially at local level such as counties, since counties are just small regions and are more sensitive to exogenous factors.

Some other challenging factors involve country size, territorial location, lifestyle and development of a country. "*Analysts tend to work more with larger countries*", however the accuracy of large counties' projections are likely to be higher (Dotson, 2018). In less-developed countries fertility rate is the most influential parameter and assumption of the future population size since fertility levels are usually high. "*Years of high fertility produce a young population age structure, which generates momentum for future growth as these youth begin having their own families*" (PRB, 2001). However, "*less-developed countries tend to have less reliable birth and death rate data*" due to epidemics and diseases, high levels of infant mortality and poor quality of statistical estimates (Dotson, 2018). In addition, "*with climate change, political unrest and any other unforeseen events, migration patterns could change unexpectedly*" (Dotson, 2018). As a consequence of all these challenges and exogenous factors the accuracy of population projections goes down.

Except for parameters, projections are also based on approximated scenarios of how the future might unfold. Probably the most common approach of creating scenarios is to elaborate together with a medium (baseline) variant a low variant and a high variant of the same projection. "*The net population increase or decrease over the period is added to the baseline population to project future population*" (PRB, 2001), and after that higher or lower vital rates are assumed at the base period. However, according to O'Neill et al. (2001), the UN variants consider various fertility tracks, but do not do so with mortality and migration.

Beside this very common way to present different variants, Caswell & Gassen (2015) have defined three other major approaches of creating scenarios to be used separately or jointly by demographers in projection construction.

1. Extrapolation of trends.
   The main idea is that the current vital trends are observed, and by assuming that they will continue to develop in this way gradually over the time, those patterns are extrapolated into the future. "*The best-known of these is perhaps the Lee-Carter model for mortality, which projects mortality with a time-series model applied to a singular value decomposition of a past record of age- and time-specific mortality rates*" (Caswell & Gassen, 2015).

2. Assumptions and expert opinion.
   This approach involves that the future trends of vital rates are simply assumed, based on unspecified conceptual models. "*The projections of Eurozone countries by Eurostat, for example, are based on the assumption that the mortality and fertility of all European*

countries will converge to a common value by the year 2150. The rates for a given country in each year are determined by interpolating between the rates at the start of the projection and the final target rates" (Dotson, 2018). In some cases, demographers require the expert opinion, which is usually called conditional, and base the future trends on the opinion of experts that are not directly involved in the projectioning process. For example, Wolfgang Lutz, an Austrian demographer and specialist in demographic analysis and population projection, used "*a Delphi-method based approach to collect and aggregate external expert opinions on demographic trends in a systematic manner*" (Caswell & Gassen, 2015). In addition, expectations of households about their own lives can be also taken into account to draw the scenarios. For instance, it could be surveyed on the expected remaining life expectancy or expected number of children.

3. Dependence on external factors, which can be projected themselves.
   The vital rates can be influenced by some external factors. If the trends of these external factors can be somehow predicted, this prediction can serve as the base for the projected vital rates. Especially this approach is widely used for the populations of animals. For instance, "*projections of populations of polar bears and emperor penguins under the impact of climate change have been based on projections of sea ice conditions generated by models of global climate conditions produced by the Intergovernmental Panel on Climate Change*" (Caswell & Gassen, 2015). In the same way, human population projections can be based on the expectations in the dynamics of economic, social or environmental developments.

4.2.1 Fertility

"*In population projection, it is necessary to anticipate the number of persons who will be born and will survive to replace the present generation*" (American Planning Association [APA], 1950). Usually, crude birth rate (number of live births per 1000 persons in population) is used. The extension of the crude birth rate is age-specific fertility rates (the number of births per each 1000 women of the ages 15–49). The age-specific fertilities are transformed into the total fertility rate of a country, which is "*the average total number of children a woman would have given current birth rates*" during her child-bearing ages (PRB, 2001).

In most developing regions, the total fertility rate is still above the level of 2.1 that represents the requirement of the precise replacement. This explains the growth of population sizes of the developing countries. However, "*the UN has assumed that their fertility rates will decline to replacement level and remain constant thereafter*" (PRB, 2001). In the most developed countries, there is a tendency of giving a birth less than to two children, and "*experts have been engaged in a spirited debate about whether fertility will continue to fall, level off, or rise again to stabilize at replacement level*" (PRB, 2001). "*If this low level of childbearing is maintained in future decades, declines in population size will occur unless the deficit in natural growth is offset with a flow of immigrants*" (NRC, 2000).

In fact, the age-specific fertility rates are more useful for projection purposes rather than the crude birth rate, since birth rates vary from woman to woman. "*The procedure which may be used for projection purposes requires the input data about: number of births, age of mothers, and number of married women of child-bearing age; that is available from census data and vital statistics* *data*" (APA, 1950). The further step is to determine what the expectations about these birth rate trends in the future are.

4.2.2 Mortality

For measurement of mortality in a country, usually crude death rate is used (the number of deaths per 1000 persons in population). The refinement of the crude death rate is age-specific mortality rates that are expressed as a proportion of deaths of a particular age to total population of this age. The age specific-mortalities tell us the information about what is the probability to die at a particular age for the given population. "*The more refined the death rate, e.g. the more detailed information that is available on the relation of deaths to sex, age, racial, income and other characteristics, the more useful it is as a tool for projecting future population*" (APA, 1950). With the help of age-specific mortalities life expectancies at particular ages can be calculated, assuming that the age-specific mortalities will hold true throughout the whole lifetime of those individuals. Life expectancy is one of the most vital outputs of population projections, especially when a population has been facing the issue of ageing. For the last decades "*life expectancy levels have risen worldwide for a long period and are projected to continue to do so, adding somewhat to future population growth and substantially to population ageing*" (NRC, 2000).

"*Mortality rates will differ in different sections of the city. High rates are likely to be found in areas populated largely by foreign born, and low rates are likely in the suburbs which are populated by young people*" (APA, 1950). It is important that demographers carefully follow the mortality traits such that life expectancies will not be underestimated. "*While these underestimates had little effect on overall population totals, they understated the future size of elderly populations and, accordingly, the looming challenges of population aging for retirement and social security programs*" (PRB, 2001).

There are external factors, like HIV/AIDS epidemic, that can suddenly distort the projected life expectancy. The example, taken from Population Reference Bureau (2001), is saying that HIV/AIDS epidemic caused an unexpected demographic crisis by lowering the projected life expectancy for 45 countries in Saharan Africa where infection rates reached 2 percent of the whole population. Then the UN estimates proved that in the 9 most affected countries new AIDS mortality had lowered the projected to 2015 population by almost 18 percent in comparison with what it would have been without AIDS. Consequently, if HIV/AIDS infection would spread significantly over other regions of the planet, it could lower the life expectancy all over these regions and strike the growth of the world population (PRB, 2001).

4.2.3 Migration

Migration is the major margin to population change; however, it is the most challenging part of the population model. On the one hand, there is emigration which patterns can be studied and described by rates, estimated by the number of individuals at risk and further applied to the relevant components to project the future population (Caswell & Gassen, 2015). On the other hand, there is immigration, the activity that cannot be characterized as the population at risk, and consequently, cannot be described as rates (Caswell & Gassen, 2015).

There is no big concern about the movements of people inside countries. The world population growth is not affected by the movements of people between countries. However, on the local level it is more important to follow the in- and out-flows since they contribute to the national growth a lot.

"*Future international migration is more difficult to project than fertility or mortality because migration flows often result from short-term changes in economic, social, or political factors that are hard to predict or quantify*" (PRB, 2001). There are no exact methods to predict sudden massive migration flows. The best thing that can help to demographers is to check the newest available estimates of the population.

There are several reasons that may trigger the movement, and one of those was World War II. According to American Planning Association (1950), "*19.5 million persons made major moves during*". However, the presence of job is one of the key causes to migrate. In the past, "*one of the major causes of the movement from farm to city has been the mechanization of agriculture, the few jobs on farms, and the lack of other job opportunities in rural communities*" (APA, 1950). These workers after were returning back to the farms having become unemployed during the depression. However, this pattern is unlikely to repeat nowadays and in future.

Demographers must assess the employment situation in the world in order to be able to make the assumptions about the future migration flows. If there are a lot of job opportunities in a country, the immigration can be expected. And if unemployment is presented in an area, the eventual emigration will be inherent to this area. However, it is easy to overestimate employment mobility, since there are so many reasons that stand behind it and a lot of things that have to be considered before the movement, like money, social attachments and family. In addition, migration factors are not all about economy. People might seek better living conditions, better climate and environment, closer culture mentality or family reunion.

4.3 Sensitivity of Parameters

The results of future projections jointly depend on the assumed levels of fertility, mortality and migration. Each of these parameters contributes differently to the total population growth. Another question is that how some changes in the assumed parameters will influence the final output of the projection. To be able to determine sensitivity of the results to the main parameters, it is needed to understand which population we are working with. The behavior of the sensitivities depends on the structure of a population and the cohorts that comprise it (Caswell & Gassen, 2015). On the one hand, if we are working with a big population any changes in total fertility rate will affect the resulting population size a lot. On the other hand, if a country is a favorable place to move in, and migration considerably contribute to this population, then the parameter of net migration must be carefully treated. If a projection with different variants is made on the basis that just one parameter's rates are varying across variants providing other parameter's assumptions stay fixed, the output of all the variants will transparently demonstrate how sensitive the resulting projection to the changes of this particular parameter is. If a population is rather small, then likely the changes in the parameters will cause just slight changes to the results.

Measuring the sensitivity of the projected population to the parameters is out of the scope of this study, however the author of this work finds it interesting to mention the way how the sensitivity can be estimated since it outflows from how the projection's assumptions are set. There is a special tool that is called Sensitivity analysis or perturbation analysis. It provides the information about "*how the results of the projection would change in response to changes in the parameters*" (Caswell & Gassen, 2015).

According to Caswell & Gassen (2015) there are several advantages of using such an analysis:

1. It can project the consequences of changes in the vital rates.
2. It can be used to compare potential policy interventions and identify interventions that would have particularly large effects.
3. It can be used retrospectively to decompose observed changes in an outcome into contributions from changes in each of the parameters
4. It can be used to identify parameters the estimation of which deserves extra attention, because they have large effects on the results.
5. It can quantify uncertainty of projection results: given the uncertainty in some parameter θ, and the sensitivity of an outcome of interest to changes in θ, it is possible to approximate the resulting uncertainty in the outcome.

Basically, there are two measures that can judge the impact on the results caused by larger or smaller changes in parameters: sensitivity and elasticity (Caswell & Gassen, 2015). They are based on the calculation of derivatives of the projected data to the parameters of initial data. Assuming that *y* is a function of *x* we define that:

$$sensitivity = \frac{dy}{dx}.$$

Sensitivity is calculated with the help of differentials and says what the sensitivity of dependent variable to changes in influence variable is. The result bears the information about how sensitive the output is to a parameter. For example, by how much the final population size will increase, if we increase fertility rate by one unit.

$$elasticity = \frac{xdy}{ydx} = \frac{\partial y}{\partial x}.$$

Elasticity is calculated with the help of derivatives and says what the proportional sensitivity of dependent variable to influence variable is. The result bears the information about what is the proportional change in the output resulting from proportional change in a parameter. For example, by what percent the final population size will increase, if we increase net migration by 5%.

4.4 Methods to project population

At the primitive level population projection can be classified on the basis of "direct method projectioning" which considers total population as a quantity that itself changes; and on the basis of "separate factor method projectioning" which involves that total population is broken down into births, deaths and migration (Renkou, 1980). No matter which classification a projection belongs to, there are much more complex mechanisms that stand behind it. In literature we can find many concrete methods which calculate future population sizes. According to Webster (2011) they are divided into simplistic (mathematical), econometric and microsimulation. These methods are comprised in the table 2.

Table 2 Alternative methods of projectioning

Class	Method	Description
Simplistic: 1. Extrapolation or projections of a trend using historical data 2. Constant increments, constant percentage change	Arithmetic increase method	The rate of growth is assumed to be constant. This method gives too low estimate and can be adopted for projecting populations of large cities which have achieved saturation conditions.
	Uniform percentage of increase	The rate of growth is assumed to be uniform rate and proportional to population.
	Logistic method	This method involves two growth rates, one of them is geometric growth rate for low population, and another is declining growth rate as country approaches some limiting population. As a result, the method has S-shape graph.
	Declining growth method	The method assumes that country has some limiting saturation population. Thus, the growth rate is a function of population deficit.
	Curvilinear method	This method involves the graphical projection of the past population growth curve, continuing according to trends based on historical data.
	Incremental increase method	The rate of growth is assumed to be progressive increasing or decreasing rate rather than constant. The average of net incremental increase of every future decade is added to the regular growth rate.
	Geometric increase method	The percentage of increase in population from decade to decade is assumed to be constant. This method provides high output and can be useful for populations with unlimited expansion.
Econometric	Time series	Population projections are based on the analyses of time series of either aggregate population size, or of vital rates. Future population size is broken down into regular components of time series: trend, seasonal, cycle and residual factor. The time series are based on historical data. Such methods may in fact be very accurate over short time horizons.
	Multiple regression	Population is estimated as a function of Instrumental Variables. Several models are estimated with the help of least squares estimators and the best model is selected.
	Econometric models	Population is predicted with a system of simultaneously interdependent equations. These models are preferred by economist because of their theoretical soundness. However, they are not much more accurate in practice than multiple regression models.

Microsimulation	Microsimulation Models	Microsimulation treats each individual independently and uses repeated random experiments instead of average probabilities. This technique simulates life events for each individual, results are then scaled to the size of the total population. It is usually based on sample instead of whole population to mitigate the complexity of calculations.

Source: Webster, 2011; O'Neill et al., 2001

4.4.1 Cohort-component method

Although the methods mentioned in the table 2 have always a place to be in demography, especially speaking about individual countries or regions, O'Neill at al. (2001) have classified those methods as alternative techniques. The most common way to project future population is called cohort-component method which is the standard method used by the national statistical agencies in the most advanced countries and global statistical offices for producing the world projections. The cohort-component method has become the prevailing way of projectioning for several reasons. First, it is not rigid, thus it can be adjusted in many different ways to suit the data whilst keeping its underlying logic (Planning Tank, 2017). Second, its fundamental trait is that "*the projected size and age structure of the population at any point in the future depends entirely on the size and age structure at the beginning of the period and the age-specific fertility, mortality, and migration rates over the projection period*" (O'Neill et al., 2001). Also, the vital rates are usually based on expert opinion.

The general idea is that initial population are divided into cohorts by age and sex, and each age- and sex-specific group are treated according to the assumptions about fertility, mortality and migration. The age-specific mortalities define the survival of each cohort forward to the next age group separately for males and females (O'Neill at al., 2001). The age-specific fertilities are applied to each female cohort during the childbearing ages (15-49 years) to calculate the total number of births which is further divided into males and females by assumed sex-ratio (Planning Tank, 2017). Similarly, age- and sex-specific net migration rates are applied to each cohort providing that "*immigration equals emigration when summed over all regions*" (O'Neill at al., 2001).

Commonly, five-year age groups with five-year time step are used for long time horizon projections. This can possibly reduce the accuracy, however, save the time. The example taken from "A Guide to Global Population Projections" (O'Neill at al., 2001) reports that "*the number of females in a particular population aged 20-25 in 2005 is calculated as the number of females aged 15-20 in 2000 multiplied by the assumed probability of survival for females of that age over the time period 2000-2005*". Such a sequence repeated separately for males and females until the projected data is reached. Of course, for this simplified calculation slightly different data is needed: age-specific vital rates must be expressed with five-year step.

According to World Population Prospects [WPP] (2015), cohort-component method cannot be considered as a complete independent projection method, since it needs basic parameters (fertility, mortality, migration) to be projected in advance as an input data for the cohort-component procedure. "*Rather, it is an application of matrix algebra that enables demographers to calculate the effect of assumed future patterns of fertility, mortality, and migration on a population at some given point in the future*" (WPP, 2015).

4.5 Uncertainty of population projections

"*Projections are inevitably uncertain*" (NRC, 2000).

As was pointed out by many global authorities there is always a place to be for uncertainty when we speak about population projections. Keyfitz (1982) says that "*Projection error exists because our understanding of demographic behavior is not perfect*" and this is absolutely true. "*The present demographic situation is not known perfectly, and future trends in births, deaths, and net migrants are subject to unpredictable influences*" (NRC, 2000). In addition, there are many other social, economic, technological, political, environmental and scientific changes, along with government policies that are influencing current and future demographic trends and population growth. Fertility and migration are most of the action. Government policies, which involves public health services, family planning methods, immigration regulations, social policy, not only affect the demographic trends but they themselves can be caused by consideration of population projections. This all together complicates the process of constructing accurate projections because "*assumptions about the future might be outmoded or invalidated in a rapidly changing industrial society*" (APA, 1950).

The span of future population projected by different global authorities is so huge because each of them relies on their own assumptions, predict the changes of environment in their own manner and use own methodology. For example, "*the U.S. Census Bureau predicts world population at 9.1 billion in 2050, compared with 9.3 billion for the latest medium projection by the UN, 8.7 billion from the World Bank, and 8.8 billion from IIASA. By 2100, the differences in the central estimates of these institutions widen to a billion or more, and differences between the low and high scenarios span more than 10 billion – from 4 billion to 16 billion*" (PRB, 2001). The variety of projected results itself causes more and more uncertainty of population projections among their users.

"*The important limitation of accuracy studies is the short history of projections (about 50 years) compared to the time horizon of future projections (100-150 years)*" (O'Neill, Balk, Brickman & Ezra, 2001). There is no way to directly evaluate the current population projections; however, the accuracy of past projections can be examined. Generally, population projections tend to be less accurate under particular circumstances. The accuracy of projected data depends on "*the quality of the input data and the assumptions made about the course of future change*" that are in fact are the most important source of errors (Environmental Systems Research Institute [ESRI], 2007). According to PRB (2001), O'Neill et al. (2001) and Bull (1987) the accuracy is lower:

1. In developing countries than in developed countries partly due to limitations and lower reliability of statistical data for their populations.
2. For smaller territories than for larger ones partly because demographers treat larger countries more thoroughly and partly because projections of smaller territories have higher error margin.
3. For younger and older age groups than for middle age groups since mistakes implied in the assumptions about mortality and fertility have greater influence at those ages.
4. At country level than at regional or global level because the error may not be balanced by some influential factors. Countries are more sensitive to errors coming from unforeseen events or migration assumptions and these errors partly cancel each other when projections are aggregated to big regions and to the world.
5. For long-term projections (more than two decades) than short-term due to the compounding effects of incorrect assumptions over time.
6. For populations with low life expectancy and high fertility due to the higher margin of error.

In projectoning, the most important sources of error are the bad quality of population estimates and collection of input data. It is needed to distinguish whether inaccuracy results from errors in assumed vital rates or from errors in baseline data.

There is no particular technique that can improve accuracy of projections. "*The key to accurate projectioning lies in evaluating your data sources and applying the information to develop reliable projections*" (ESRI, 2007). "*Demographers try to measure the uncertainty of population projections by consulting other experts; analyzing errors in previous projections; and examining trends in fertility, mortality, and migration*" (PRB, 2001).

"*Recent projection methodologies have focused on identifying uncertainty in projections*" and it is recommended for demographers to "*develop new ways to characterize the uncertainty that is associated with any population projection*" (PRB, 2001). One of them is to state the probability that future population size will fall into particular region. These methodological refinements together with continued improvements in the assumptions and evaluation of the previous projections can make population projections more credible and useful for a wider range of users.

5 METHODOLOGY OF ACCURACY EVALUATION

As was mentioned in the previous chapter there is no direct method to evaluate the errors of current projections, however it is possible to post evaluate the accuracy of projections. According to ESRI (2007), there are several approaches to evaluate the performance of demographic projections:

- Examining projection compared to historical patterns of population change
- Comparing projection to other estimates or projections for the projection area
- Submitting projection to knowledgeable persons in the projection areas for assessment
- Performing sensitivity analyses by testing the effects of different methods and assumptions
- Comparing projection with known population values such as a census count

Different studies demonstrate that the most common way is to conduct the evaluation on the basis of comparing projection to other population estimates, the real values that are observed by national and global statistical agencies. For instance, ESRI (2007) and Smith & Shahidullah (1995) in their papers use the Mean Absolute Percent Error [MAPE] and the Mean Algebraic Percent Error [MALPE]. Swanson, Tayman & Bryan (2011) recall in their study the Mean Square Percentage Error [MSPE], the Root Mean Square Percentage Error [RMSPE], the Mean Absolute Percentage Error [MAPE] and the Median Absolute Percentage Error [MEDAPE]. Also, in studies of Šídlo & Tesárková (2009) and Šmejkalová (2011) we can find the Keyfitz's "Quality of Prediction Index". Pflaumer (1992) suggest that the evaluation should be based not on the comparison of projected figures and observed values but on the comparison of actual average annual and projected average annual growth rate.

In fact, the outcome of all the techniques mentioned above should be approximately the same, and all of them serve the similar logic. As a result, we get the percentage that bear the information about how far the projected data is deviated from the actual data either for a concrete year and component, or for the whole period on average. "*Percent errors were used to identify outliers, or extreme differences; and average percent error summarizes the relative differences by geography*" (ESRI, 2007). However, the question is which errors can be considered as low or high. There is no a particular criterion or range of better or worse of accuracy that can be found in literature. The thing is that criteria must be set individually in accordance with the specification of a given population, since "*magnitude of the error is inversely related to the size of the population – the smaller the population, the larger the error*" (ESRI, 2007). The percentage error can be very close to zero which indicates excellent performance and high accuracy of a projection; and it can reach very high percentages and even exceeds 100% for some of the components and parameters, which indicates that the accuracy goes down. ESRI (2007) states in its paper that for the projected total population size at the national level an error of 5% or more is considered high while an error of 10% or more is considered high at the regional level. According to another study, the distribution of errors is that everything what is below 10% can be considered as small errors, and everything what is above 25% can be considered as large errors (Smith & Shahidullah, 1995).

For measuring the accuracy of the individual projections and their parameters two quantitative methods, namely the Keyfitz's "Quality of prediction index" [Keyfitz] and the Mean Absolute Percentage Error [MAPE], are used in the work. The choice for Keyfitz is justified because of its simple use and separate employment to any of the components and parameters of projections. MAPE is widely used and accepted by many users of population projections, which makes its results comparable to many other studies worldwide.

5.1 Mean Absolute Percentage Error

MAPE is the most commonly used method to evaluate the accuracy of population projections. As evidence, besides its common usage and reference in various demographic studies and population papers, it also can be often found in software packages such as Autobox, ezProjectioner, Nostradamus, SAS, and SmartProjection. (Swanson, Tayman & Bryan, 2011). There are several advantages of using MAPE and according to Swanson et al. (2011) they are:

- valuable statistical properties such that it makes use of all observations and has the smallest variability from sample to sample,
- clarity of presentation because it is expressed in generic percentage terms that is understandable to a wide range of users that makes it useful for purposes of reporting,
- absolute terms that do not let negative and positive values cancel each other.

MAPE is popular due to simplicity and ease to understand it, however there is one big disadvantage of this method. MAPE, like any other average, is sensitive to extreme value, and it must be relevantly chosen. In addition, MAPE does not show the direction of error, and it is not possible to determine whether projection is over- or underestimated. MAPE in some cases can be transformed to MAPE-R (re-scaled), which eliminates the outliers in the series of projections, as the outliers have a big impact on the final calculation.

MAPE is usually compared to the Mean Percentage Error [MPE] or is given as an improved version of MPE. MPE judges the accuracy of projection by the computing the percentage deviations of projected values from the actual data both in underestimating and overestimating ways (Clements, 2016).

Formula:

$$MPE = \frac{1}{n} * \sum \frac{R(t) - P(t)}{R(t)} * 100.$$

MAPE eliminates negative values when calculating deviation that provides the whole idea about what is the average percentage error in multiple projections. Since negative and positive values exclude each other, we cannot see the real average error. However, MAPE does not show whether projection is overestimated or underestimated against the actual data (Clements, 2016). In this work MAPE method is presented for calculating the coefficients.

Formula:

$$MAPE = \frac{1}{n} * \sum \frac{|R(t) - P(t)|}{R(t)} * 100.$$

The example below shows the difference between MPE and MAPE (data taken from the CSU-2009 projection, medium variant, total population 2010):

Table 3 Comparison of MPE and MAPE

	Projection	**Reality**	+/-	**MPE**	**MAPE**
2009	10 467 542	10 467 542	0	0	0
2010	10 503 408	10 506 813	3405	0,0003241	0,0003241
2011	10 538 186	10 486 731	-51455	-0,0049067	0,0049067
2012	10 572 374	10 505 445	-66929	-0,0063709	0,0063709
2013	10 605 542	10 516 125	-89417	-0,0085028	0,0085028
2014	10 637 468	10 512 419	-125049	-0,01189536	0,01189536
2015	10 667 999	10 538 275	-129724	-0,0123098	0,0123098
2016	10 697 033	10 553 843	-143190	-0,0135676	0,0135676
2017	10 724 526	10 578 820	-145706	-0,0137734	0,0137734
Total	95 414 078	94 666 013	Average	**-0,00788916**	**0,008956322**

Source: own calculation

5.2 Keyfitz's "Quality of prediction index"

The Keyfitz's index is one of the simplest ways to evaluate the quality of a projection. This coefficient judges the quality against the "standard" if there is so. The so-called standard value could be: 1) default population or extrapolation, 2) projected values based on assumptions, 3) or it simply could be equal to zero (in case we deal not with forecast but with projection itself) (Šídlo & Tesárkova, 2009). The last case demonstrates the result as the usual fraction with projected value at the numerator and actual value at the denominator.

Formula:

$$Q(b,t) = \frac{P(t)-b}{R(t)-b}, \text{ where}$$

Q(*b*,*t*) – quality index of the chosen projection for a particular time *t* and standard *b*
P(*t*) – projected value at time *t*
R(*t*) – actual value at time *t*
b – standard value

Due to the fact that the projections were released in the different calendar years, it is more appropriate for the comparison to refer time as the time elapsed after the projections' release rather than the individual calendar years because the last consider the actual development regardless the time spent after their publications (Šídlo & Tesárkova, 2009). This can be also supported by the fact that with time spent accuracy of projections goes down, and newest projections may seem more accurate than old ones. In this work the Keyfitz's index with the zero standard value is used because there is no such standard value that the projections (they are not forecasts) can be compared.

5.3 Keyfitz's "Quality of prediction index" by weighted age groups

Following the previous sub-chapter, the accuracy of population projections depends not only on the time spent after their releases but also on the age groups since the number of persons in each group is determined by individual factors and assumptions (Šídlo & Tesárkova, 2009). For that reason, Keyfitz weighted by age groups is introduced. This is the expanded version of the simple Keyfitz's coefficient (5.2) with the zero value of standard and weighted by age groups. It enables to include the simple Keyfitz's coefficients for the age groups we are interested in and weight them into the total composite index.

Formula:

$$Q_a(0,t) = \frac{\sum_{x=0}^{\omega} W_x |1 - Q_x(0,t)|}{\sum_0^{\omega} W_x}, \text{ where}$$

x – age specific category
W_x – weights for age category x using the real values, weight is equal to population of the particular age group divided by the total population
$Q_x(0,t) = P_x(t)/R_x(t)$ – quality index of the projection for a particular age category x,

The disadvantage of this indicator is that the positive and the negative age variations can be mutually cancelled which leads to the underestimation of the final average index (Šídlo & Tesárkova, 2009). The index of each single age group is calculated in absolute value such that the index expresses the deviation regardless its direction. If the Keyfitz's index weighted by age groups is calculated for all age groups, thus the sum of all weights, the denominator in fact, is equal to 1 (Šídlo & Tesárkova, 2009).

The example below shows the procedure of calculating Keyfitz weighted by age groups (data taken from the medium variant of CSU-2009 males, 2010):

Table 4 Example of Keyfitz's index

Age categories	Age coefficients	Age weights
0-14 years	$Q_{0\text{-}14} = 0.9993$	$W_{0\text{-}14} = 0.1493$
15-64 years	$Q_{15\text{-}64} = 1.0006$	$W_{15\text{-}64} = 0.73$
65+ yeas	$Q_{65+} = 1.0012$	$W_{65+} = 0.1247$
$\boldsymbol{Q_{0-14,15-64,65+}(0,2010)} = \frac{((0.1493*\lvert 1-0.9993\rvert)+(0.73*\lvert 1-1.0006\rvert)+(0.1247*\lvert 1-1.0012\rvert))}{0.1493+0.73+0.1247} = \mathbf{0.0006975}$		

Source: own calculation

5.4 Interpretation of the results

Based on the written literature and the principle of understanding the distribution of lower or larger errors that differs from population to population, the boundaries of better or worse accuracy were determined by the author of this work and are presented in the table below. These boundaries respect the specification of the population of the Czech Republic and are set up according to the outcome of the calculations done the projections. For example, if the deviation of the total population for the most variants does not exceed 1%, and just few of them belongs to the interval [0%; 0.1%], and few of them exceed the value of 1%, that means that the boundary for the good accuracy is less or equal to 0.1% and the boundary for the bad accuracy is more or equal to 1%.

So, for example, if the boundary set for the good accuracy is <0.1%, and the boundary for the bad accuracy is >1% that means that the interval that falls between good and bad we can consider as middle accuracy, in our case it is (0.1%;1%). The same logic will be for the age category 100+ years but with higher intervals because of the difficultness of predicting this age group. So, we have the boundary <10% for the good accuracy, the boundary >50% for the bad accuracy, consequently for the middle accuracy we have the strict interval from 10% to 50%.

For the interpretation of the results the following rules should be addressed. All criteria can be found in the table 5.

Table 5 Boundaries of accuracy for Keyfitz and MAPE

Component	**Good accuracy**	**Bad accuracy**
Total populations	Deviation <= 0.1%	Deviation >= 1%
Age categories 0-89 years	Deviation <= 0.1%	Deviation >= 1%
Age category 90-99 years	Deviation <= 1%	Deviation >= 5%
Age category 100+ years	Deviation <= 10%	Deviation >= 50%
Total fertility rate	Deviation <= 1%	Deviation >= 5%
Median age	Deviation <= 0.1%	Deviation >= 1%
Net migration	Deviation <= 10%	Deviation >= 50%
Life expectancies at birth, at 65 years	Deviation <= 0.1%	Deviation >= 1%
Crude death rate	Deviation <= 1%	Deviation >= 5%
The rest of the components	Deviation <= 0.1%	Deviation >= 1%

Source: own elaboration

6 CURRENT DEMOGRAPHIC TRENDS IN THE CZECH REPUBLIC

Population of the Czech Republic has been constantly increasing since 2003, except for the year 2013 where both negative balance of natural change and net migration were traced (Křesťanová, Kurkin & Šafusová, 2018).

At the beginning of 2017 the population size of the Czech Republic accounted for 10 578 820 persons out of which there were 5 200 687 males and 5 378 133 females which evidences a little prevalence of women oven men (Czech Statistical Office [CZSO], 2018). The mean age of the population was 42 years. The portion of people aged 65+ was 18.8% of total population, which is about 1.9 million people, that can be defined as aged society (more than 14%) and almost reaching super aged society (more than 21%) (Miskolczi & Cséfalvaiová, 2013), and we can observe the constant trend of increasing the number of people aged 65+ since 1920.

By the end of 2017 the population size has exceeded the threshold of 10.6 million accounted for 10 610 055 inhabitants on 31.12.2017. The total increase of the population was 31.2 thousand people during the year 2017, and international migration was the dominant source of increase and equalled to 28.3 (CZSO, 2018). The natural increase contributed a little to the total population adding just 2962 persons (CZSO, 2018). The population increases significantly in the senior part of the population. The population of the Czech Republic is ageing, and it started back in the 1980s. This process is reflected in the increasing mean age of the population, the median age, and the index of ageing.

For the first half of 2018, the population of the Czech Republic has increased by 15.4 thousand primarily thanks to net migration that balance was estimated as 17.7 thousand (CZSO, 2018). The natural change resulted to be negative and decreased the population size by 2.3 thousand (CZSO, 2018).

6.1 Fertility

The total fertility rate in the Czech Republic has been raised since 2011 where it was equal to 1.43, and at the end of 2017 it equalled to 1.69 making it the highest one since 1992 (Křesťanová et al., 2018). This rate is above the average European fertility which equals to 1.6 for EU-28. The most fertile ages are fall into the age group 30-34 since 2011, and the mean age of mothers for giving a birth has been equaled to 30 in 2017 (Křesťanová et al., 2018). There is a steady trend of the rising total fertility for the last years.

According to CZSO (2018) there were 114 405 live births recorded in the Czech Republic, which is by 1742 more than the year before. The year 2017 brought the highest number of births since 2011 when the number of newborns accounted for 117 153 (Křesťanová et al., 2018). In the first half of 2018, 55.7 thousand children were born which is by 0.5 thousand less comparing to the previous year. Similar to 2017, the most children were born to women aged 30 while the first child was born to females aged 27 (CZSO, 2018).

6.2 Mortality

During 2017, 111 443 people died in the Czech Republic which indicates the increase in the number of deaths by 3693 comparing to 2016. This is the highest death number for the last 20 years in the history of the Czech Republic that was caused by the widely spread epidemic of influenza taken away about 10-12 thousand per month (CZSO, 2018). The portion of men died was by little more than the portion of women (51% against 49%), and the most frequent ages to die were 70-74 for the male part and 85-89 for the female part respectively (CZSO, 2018). Men life expectancy at birth was equalled 76 years, and women life expectancy at birth reached 81.9 years in 2017, and both of these parameters indicates the prolongation of lifespan in the Czech Republic over the last years (Křesťanová et al., 2018).

In the first half of 2018, there were 58.1 thousand of deaths which is very similar in comparison to the first half of 2017. Again, more men died than women: 29.2 thousand of men and 28.8 thousand of women (CZSO, 2018). "*Owing to higher mortality men than women and numerically large cohorts of births after World War II, men aged 70–74 years were predominant among deceased men*" (CZSO, 2018).

6.3 Migration

In total, 45 957 people have immigrated to the Czech Republic from abroad in 2017 which is by 8.5 thousand more than in 2016 (Křesťanová et al., 2018; CZSO, 2018). Opposite to, 17 684 people have left the country which is 0.2 emigrants more than the year before (Křesťanová et al., 2018; CZSO, 2018). Those values present the highest number of migration out- and inflow since 2009 (Křesťanová et al., 2018). The net migration was equal to 28 273, 8209 people more than in the previous 2016, and was the highest since 2010 (Křesťanová et al., 2018; CZSO, 2018).

In 2013, the balance of migration was uncommonly negative and that happened just once over the last decade. That happened because the amount of people left the country (about 31 thousand) was by 1297 persons more than the amount of people arrived in the Czech Republic (about 30 thousand) (International Migration Outlook [IMO], 2015). According to IMO (2015), the migration outflow almost reached a pre-crisis level (33 thousand people) that was registered in 2006. However, the number of immigrants remained far below pre-crisis levels.

There were more males who immigrated from abroad and emigrated from the Czech Republic. "*Males made up 58.4% of immigrants and 56.3% of emigrants in 2017*" (Křesťanová et al., 2018). Positive net migration over the last years was composed mainly from the age groups from 15 to 34 years (Křesťanová et al., 2018). By five-year age groups the highest net migration was observed in the age group of 25-29 years.

In 2017, the population of the Czech Republic has grown significantly thanks to the positive net migration which was made up of nationals of Ukraine (7690), Slovakia (4356), Romania (1602), Bulgaria (1437), Russia (1346) and Vietnam (1316) constituting together 63% of the total net migration (Křesťanová, Kurkin & Šafusová, 2018).

During the first half of 2018, the population of the Czech Republic has grown by 17.7 thousand people due to the balance of migration which is by 5.2 thousand persons higher than the first half of 2017 (CZSO, 2018). The number of emigrants and immigrants are increasing comparing to the previous years.

7 POPULATION PROJECTIONS IN THE CZECH REPUBLIC

As was outlined in the chapter 6, the population development of the Czech Republic has experienced many significant changes and continues to do so. The unexpected drop in the net migration during 2013 followed by continues increase of migration inflow and outflow as well as the net migration as well. Also, there is a steady trend of population ageing in the country while increasing life expectancy at birth. In addition, we can see that the population itself is growing together with total fertility rate that is going up. That all together makes it interesting to answer the question how the future population will look like.

The changes in the demographic conditions influence all spheres of life. That explains why population projections are becoming more and more interesting for the demographers and policymakers. Consequently, the main question arises whether the projections of the Czech Republic's population are accurate enough and how they can be better used to meet political, economic or social reforms.

This chapter covers the information about the current population projections of the Czech Republic's population. Also, the results of the previous evaluation of the past projections, that has been done just once, are discussed.

7.1 Current population projections

There are several official authorities that are constantly issuing the projections of the population in the Czech Republic. The Czech Statistical Office [CZSO] is the national statistical agency of the Czech Republic. Eurostat is the European Statistical Office whose main responsibilities are the provision of statistical information to the institutions of the European Union, and advisory of the harmonized methodology for statistical processes across its member states. The statistical division of the United Nations is the global incorporation that serves the statistical needs and coordinates the activities of the global statistical system. There world population projections are called World Population Prospects [WPP]. Also, there are two individual researches in the Czech Republic being the professors of the Charles University, department of demography, who produce population projections for the private usage. Their names are Boris Burcin and Tomáš Kučera [B&K].

The methodological aspects of the projections of the authorities mentioned above are discussed more in depth in the chapter "Data".

7.2 Previous evaluation of population projections

As was mentioned above, the accuracy evaluation of the population projections of the Czech Republic was conducted just once by Luděk Šídlo, the postgraduate and since 2007 representative of the Faculty of Natural Science belonging to the Charles University; and Klára Tesárková, the former PhD student of the same faculty and university where she operates as an assistant since 2008 (Šídlo & Tesárkova, 2009). The evaluation was published in 2009 within the framework of the journal "Demografie", the only demographic journal published in the Czech Republic, which comprises the review for population research. Šídlo and Tesárková describe the basic evaluation methods and introduce some basic results of the accuracy evaluation of the population projections.

For the accuracy evaluation of the population projections four methods were used: the Keyfitz's "Quality of Prediction Index", the average Keyfitz's "Quality of Prediction Index" (the one with weighted age groups), Theil`s index U and evaluation method based on the principle of APC models. Within the framework of this work, taking into account the methods used by the author of this thesis, just the results of the Keyfitz's index and the weighted Keyfitz's index calculated by Šídlo and Tesárková are discussed. Those results can be somehow comparable to the outcomes of the analysis presented in the chapter 9.

In their study authors applies the assessment to projected data of four different sources. They use the projections of B&K and CSU with the base year in 2003, of Eurostat with the base year in 2004, and of WPP with the base year in 2002. The projected data is compared against the actual data estimated by the Czech Statistical Office and according to the time spent after the projections' release. For the total population size the results prove that the accuracy of B&K-2003 and CSU-2003 fluctuate around 100% for the first two years after the release. The low variant of Eurostat-2004 projection starts to deviate from the reality by 1.5% during the third year after the release, and the baseline variant deviates by 2% from the reality during the following year. The accuracy of WPP-2002 is the maximum during the third year after its publication; otherwise it is over- or underestimated throughout the various periods.

The authors pay attention that the year 2007 was critical to all the projections due to its intensive demographic development both in the number of births and the rapid increase of the number of immigrants (Šídlo & Tesárkova, 2009). However, there is no such a big gap between the actual and projected data for the all projections, and the deviation lies at the level of 0.5-2.2%. The most deviated projection is Eurostat-2004 the low variant of which reaches 2.5%. The deviation for the whole period of all the projections belongs to the approximate interval [97.2%; 100.5%].

For the age group 0-4 years, which is directly influenced by the predicted number of births, the projections of CSU-2003, B&K-2003, high variant and high fertility variant of Eurostat-2004 perform relatively accurate till 2006 with the maximum deviation of 1.3%. The projection WPP-2002 together with the low and the baseline variant of Eurostat-2004 perform worst of all and in three years after the release have the deviation of 3% and more. At the latest available year of the evaluation (2007) the projections underestimate the reality up to 10% (low variants CSU-2003 and B&K-2003, baseline variant Eurostat-2004, WPP-2002) and even up to 20% (low variant Eurostat-2004). The deviation for the whole period of all the projections belongs to the approximate interval [81%; 102%].

Another key age group that was taken under the assessment by Šídlo and Tesárková is 85 and above (for Eurostat 80+), which is influenced most of all by mortality rates, high error margin and in long-term perspective by migration (Šídlo & Tesárkova, 2009). On average, the projections perform very well except for WPP-2002 which deviation increases almost to 8% after the publication and then it rapidly goes under the 100% line and reaches 12% in 2007. Eurostat-2004 with the high, baseline and low fertility variants is not far from the actual (about 2.5%), and juts low variant underestimates the reality more (about 6.5%). Also, all the variants of CSU-2003 and B&K-2003 fluctuates around 100% line with the maximum deviation of 5-6%. The deviation for the whole period of all the projections belongs to the approximate interval [88%; 108%].

The analysis of individual age groups, where just the medium (baseline) variants were engaged into the evaluation, proved that:

- B&K-2003 projection performs with higher deviations for younger and older age groups, but also for the age group 20-24 where the deviation reaches almost 5%. Interestingly, the age group 0-4 is underestimated but the age group 5-9 is overestimated almost by the same share. The age group 85-89 is overestimated as well. The oldest age group (90+) is relatively accurate with the deviation around 5% for the latest available year
- CSU-2003 projection demonstrates similar mistakes to B&K-2003 and has worse accuracy for the younger and older age groups as well as for 20-24 years. Authors connect the inaccuracy of this group to the imperfectly set of migration assumptions. The age group 90+ is much less accurate comparing to B&K-2003 and in 2007 it exceeds the deviation of 15%
- Eurostat-2004 underestimates and overestimates different age groups throughout the all life stages. More than that, the deviations are significantly higher than the ones of CSU-2003 and B&K. The least accurate age groups are 0-4, 10-14, 30-34 and 60-64 with the deviation reaching 12%. However, age group 80+ (the oldest available) performs with excellent results throughout the whole time period fluctuating around 100%.
- WPP-2002 has the biggest deviations of projected values from the reality for all the age groups. The most extreme of them are 0-4, 80-84, 85-89 with the deviation exceeding 16% in some cases. Moreover, the same age group is overestimated in one calendar year and underestimated in another. The oldest age group 90+ underestimates the observed values by 14% in the first after the projection's release, however in the fifth year it lies almost on the 100% line, and in the latest year it overestimates it just by 4%.

On average, with support of the Keyfitz's index weighted by age groups, the domestic projections perform better than the foreign projections. CSU-2003 demonstrates the lowest average deviation, and Eurostat-2004 the highest one. It is clearly seen that the accuracy of the projections depends on the time spent after the projections' release.

8 DATA

8.1 Population projections

13 variants of 4 individual projections were used in the work, namely projections 2009 and 2013 from the Czech Statistical Office [CSU] with medium, low and high variant; baseline variant 2015 from Eurostat; medium, low and high variant produced by the United Nations in 2015 (World Population Projections [WPP]); and the projection 2009 prepared by Boris Burcin and Tomáš Kučera [B&K] with medium, low and high variant.

At the first stage CSU-2009 high variant, CSU-2013 low variant, WPP-2015 low and high variants and B&K-2009 high variant were eliminated from the evaluation, since those variants were the most significant outliers (see figure 1).

The final involvement of the projections and their variants is comprised in the table 6.

Table 6 Projections involved in the analysis

Authority	Projection	Variant
Czech Statistical Office	CSU-2009	low, medium
	CSU-2013	medium, high
Eurostat	Eurostat-2015	baseline
United Nations	WPP-2015	medium
B. Burcin & T. Kučera	B&K-2009	low, medium

The next sub-chapters describe the methodological aspects and characteristics of the projections used in this study in more detail.

8.1.1 Czech Statistical Office

CZSO issues population projections on the regular basis at the national and regional level. Among their past publications, the population projections of 2003, 2009, 2013 can be found. The last projection was issued in 2018 and it covers the time horizon till 2100. The projections of CZSO are deterministic and are processed in three variants: low, medium and high; out of which medium variant is considered as the most likely one (CZSO, 2018). Also, there is an additional variant with no migration. The main characteristics and methodology of the projections are presented below:

- projections were elaborated with cohort-component method
- reproductive behavior and mortalities by sex and age are expected to be the same among people
- the basic input data is the number of inhabitants of the Czech Republic by sex and age as of beginning of year, which comes out from the Population and Housing Census of 2011
- population is projected as of January 1st by sex and single year of age
- total number of projected live births, deaths and net migration
- projected population structure indicators: shares of various age groups in total population, age dependency ratios and mean age of population
- assumptions on age-specific fertilities rates are set for the ages with 5-year step interval
- assumptions on life expectancy are set for the ages with 5-year step interval by sex

- assumptions on net migration are set for the ages with 10-year step interval be sex

Within the framework of this study two population projections of CZSO are used: CSU-2009 with its low and medium variants and CSU-2013 with its medium and high variants. The details about the most important indicators of the projections can be found in the tables 7 and 8 below. The figures in the tables show the development of the population and the individual components as of the last year of the projections.

Table 7 CSU-2009 key figures

Projection	**CSU-2009** *time horizon 2009-2065(2066)*		
Variant	Low	Medium	High
Total population	9 053 624	10 666 055	12 391 684
Population aged 65+	2 967 203	3 411 569	3 903 184
Total fertility rate	1.55	1.72	1.85
Life expectancy at birth, males	84.1	88.6	88.5
Life expectancy at birth, females	88.5	91.0	93.0
Net migration	15 000	25 000	40 000

Source: CZSO, own elaboration

Table 8 CSU-2013 key figures

Projection	**CSU-2013** *time horizon 2009-2100(2101)*		
Variant	Low	Medium	High
Total population	6 095 234	7 683 652	9 083 414
Population aged 65+	2 038 330	2 498 583	2 947 111
Total fertility rate	1.45	1.56	1.61
Life expectancy at birth, males	84.2	86.6	88.4
Life expectancy at birth, females	88.8	91.1	92.9
Net migration	10 350	17 671	25 400

Source: CZSO, own elaboration

According to the two tables we can observe that the population size of the Czech Republic will have risen to 10 666 055 inhabitants by the beginning of 2066, however it will drop below 8 million at the beginning of the 22nd century. The same trend is observed for the several components: in 2065 we see the higher numbers than today's figures, and in 2100 the figures are above the today's reality. This is true for the total population size, total fertility rate (1.72 in 2065 against 1.56 in 2100) and net migration (25 000 in 2065 against 17 671 in 2100). Life expectancy at birth will continue growing (88.6 and 91 in 2065 against 86.6 and 91.1 in 2100) except for the male part in 2100 where the number of people aged 65+ in 2100 will be lower comparing to 2065 but still higher than now.

8.1.2 Eurostat

Eurostat produces population projections for the European countries at the national level. The past projection was issued in 2008, and the freshest one in 2015 covering 29 countries (28-member state + Norway). Due to the rapid ageing of the European population and the require from Economic and Financial Affairs Council configuration in 2015, Eurostat started to coordinate work, together with national statistical agencies aimed to develop models for calculating future populations (Eurostat Commission, 2017). During several meetings that were hold by Working Group on Population Projections chaired by Eurostat in March 2015 some methodological proposals were intensively discussed, specifically the agreement was reached concerning mortality and fertility models (Eurostat Commission, 2017).

Eurostat notifies that its new 2015-based projection should be perceived as one of the many possible future developments of European countries, and some projections created by national statistical agencies may reasonably diverge from the Eurostat's projection (Eurostat Commission, 2017). Eurostat projections are deterministic and are processed in one baseline scenario and four variants: lower fertility, lower mortality, higher migration, lower migration (Eurostat Commission, 2017). Also, there is an additional variant with no migration. In addition, the variety of such variants is aimed to conduct the sensitivity test. The main characteristics and methodology of the projections are presented below:

- projections were elaborated with cohort-component method
- the basic input data is collected by Eurostat and harmonized in accordance with specific EU regulations: complete dataset of fertility, mortality and migration data
- population is projected as of January 1st by sex and single year of age
- the main dataset of assumptions: age-specific fertility rates, age-specific mortality rates and international net migration figures
- projected population structure indicators: shares of various age groups in total population, age dependency ratios and mean age of population
- total number of projected live births, deaths and net migration
- fertility assumptions are based on the extrapolation (Autoregressive Integrated Moving Average [ARIMA] model) of four parameters: overall level of fertility, starting age of fertility, age at which fertility reaches its peak level and youngest age above peak fertility at which fertility falls to half of its peak
- the initial mortality patterns are derived from the period-cohort age- and sex-specific rates reported by the country the year prior to the projection release
- net migration is extrapolated with ARIMA model

Within the framework of this study Eurostat's projection with the base year in 2015 is used with its baseline variant. The details about the most important indicators of the projections can be found in the table 9. The figures in the table show the development of the population and the individual components as of the last year of the projections.

Table 9 Eurostat-2015 key figures

Projection	**Eurostat-2015** *time horizon 2015-2080(2081)*				
Variant	L. fertility	L. mortality	Baseline	H. migration	L. migration
Total population	7 918 574	10 045 025	9 767 193	10 311 053	9 225 530
Population aged 65+	2 771 534	3 050 684	2 792 376	2 888 450	2 696 318
Total fertility rate	1.47	1.84	1.84	1.84	1.84
Life expectancy at birth, males	86.2	88.5	86.2	86.2	86.2
Life expectancy at birth, females	90.4	92.8	90.4	90.4	90.4
Net migration	8985	8985	8985	11 966	5982

Source: Eurostat, own elaboration

According to the projection of Eurostat the total population size of the Czech Republic will reach 9 767 193 in the beginning of 2081 which is in accordance with the projections of CSU-2009 and CSU-2013. Total fertility rate (1.84) and life expectancy at birth (86.2 and 90.4) also are in accordance with the trend. However, the Eurostat's projection results in a very low net migration in 2080 which is in controversy with the projections of CSU whose predictions are much higher both in 2065 and 2080. Moreover, net migration of 8985 is quite below the today's situation.

The interesting thing that can be observed from this table is that Eurostat's projection is elaborated with such variants out of which the sensitivity of the projected population to the basic parameters can be traced, specifically lowering or rising one parameter while all other parameters fixed at the level of the baseline variant. At the first glance, we can see that lowering total fertility rate from 1.84 to 1.47 has the highest impact on the total population size (almost 2 million people) in 2080. The lowering of mortality rates resulting in 2.3 (males) and 2.4 (females) years increase in the life expectancy at birth has the least significant effect. However, it is more relevant to assess the relative changes in the output on the relative changes in the parameters, since different weights of increase or decrease in the parameter indicators will results in different changes in the total population size.

8.1.3 United Nations

United Nations projections are considered to be one of the best prepared by global statistical agencies. The last issue is the 2017 Revision of World Population Prospects is "*the twenty-fifth round of official United Nations population estimates and projections that have been prepared by the Population Division of the Department of Economic and Social Affairs of the United Nations Secretariat*" (United Nations). Their projections are produced for all countries and areas of the world. The preparation of each new Revision of the official projections involves two activities: 1) collection of new demographic information of each country in the world; 2) elaboration of assumptions about the future traits of fertility, mortality and migration for every country in the world (WPP, 2015).

Recently, certain components of the projections have started to be measured by probabilistic approach, such as total fertility and life expectancy at birth by sex (WPP, 2015). UNs projections are elaborated with a number of various scenarios mainly for the purpose to conduct the sensitivity analysis. Also, low, medium and high variants can be accessed that are extrapolated on the basis of different fertility levels. The main characteristics and methodology of the projections are presented below:

- projections are elaborated with cohort-component method since 1963 Revision
- the basic input data is estimated to be consistent with official population estimates as well as subsequent trends in fertility, mortality and international migration
- population is projected as of July 1st by sex and quinquennial population five-year age groups
- the main dataset of assumptions: age-specific fertility rates, life expectancy at birth by sex, and international net migration figures
- total fertility is based on official estimates of age-specific fertility rates
- infant and child mortality is based on the registered births and infant deaths
- life expectancy at birth is based on official estimates of life expectancy available
- mortality rates are based on life tables from the Human Mortality Database and are calculated by five-year age group and sex after determination of life expectancy at birth
- international migration is based on: 1) official estimates of net international migration 2) as the difference between overall population growth and natural increase through; also is set up according to: 1) data on labor migration flows; 2) estimates of undocumented or irregular migration; 3) data on refugee movements in recent periods
- new 2017 Revision involves the prediction intervals that provide an assessment of uncertainty

Within the scope of this study, the projection of 2015 Revision is used with its medium variant. The details about the most important indicators of the projections can be found in the table 10. The figures in the table show the development of the population and the individual components as of the last year of the projections.

Table 10 WPP-2015 key figures

Projection	**WPP-2015** *time horizon 2015-2100*		
Variant	Low	Medium	High
Total population	5 744 157	8 892 165	13 155 132
Population aged 65+	2 230 495	2 656 396	3 082 734
Total fertility rate	1.35	1.85	2.35
Life expectancy at birth, males	87.7	87.7	87.7
Life expectancy at birth, females	90.9	90.9	90.9
Net migration	30 000	30 000	30 000

Source: UN, own elaboration

As predicted by WPP, the population size of the Czech Republic will account for 8 892 165 people in the beginning of July. This number is considerably higher than the one stated by CSU-2013. The population aged 65+ and life expectancies at birth look close to CSU-2013 figures in 2100. WPP predict the highest total fertility rate and the highest net migration even by the year 2100. The low and high variants are based on the different assumptions in the total fertility rate what makes it possible to check the sensitivity of the total population size to the total fertility rate. The range of the total fertility rate is pretty wide between the low and high variants, consequently the boundaries of the resulted population are very distant from each other. In 2100, the projected population size of the Czech Republic by the low variant is the smallest among other projections and equals to 5 744 157 inhabitants as well as the projected population size by high variant is the largest among other projections and equals to 13 155 132 inhabitants with the difference of 7 410 975.

8.1.4 Burcin and Kučera

Boris Burcin and Tomáš Kučera are two individual demographers who represent the Department of Demography and Geodemography of the Faculty of Science at the Charles University in Prague and are continuously working on processing of the population projections referred for the Czech Republic. Their projections are not in free access and intended more for private purposes. The most recent known projections were issued in 2003, 2008 and 2013. The projections of B&K are processed in three variants: low, medium and high. The main characteristics and methodology of the projections are presented below (Burcin & Kučera, 2010):

- projections are elaborated with cohort-component method
- projections are elaborated in accordance with the generally accepted methodology and international recommendations
- the basic input data is the detailed, relatively reliable, and sufficiently representative statistical data
- population is projected as of December 31st by sex and single year of age
- the main dataset of assumptions: age-specific fertility rates, age-specific mortality rates or quotients, emigration rates and absolute numbers of immigrants by age
- total number of projected live births, deaths and net migration
- the projected parameters are presented as life expectancy at birth and at 65 years, total fertility rate and net migration

Within the scope of this study the population projection issued in 2009 is used with its low and medium variants. With the fact that projections are produced as of the end of year, the figures presented on December 31st are considered as the figures of the beginning of the next year (January 1st next year) for the better usage and comparison. For example, in fact, the projection starts on 31.12.2008 and finishes on 31.12.2070; however, for the purpose of this study it is taken as of 1.1.2009 with the time horizon till 1.1.2071. The details about the most important indicators of the projections can be found in the table 11. The figures in the table show the development of the population and the individual components as of the last year of the projections.

Table 11 B&K-2009 key figures

Projection	**B&K-2009** *time horizon 2009-2070(2071)*		
Variant	Low	Medium	High
Total population	9 284 882	11 101 763	12 878 715
Population aged 65+	2 667 673	3 185 997	3 643 560
Total fertility rate	1.56	1.75	1.87
Life expectancy at birth, males	81.9	84.9	86.6
Life expectancy at birth, females	86.3	88.9	90.6
Net migration	20 094	29 973	44 972

Source: B&K, own elaboration

The projection of B&K suggests that the total population size will take the number of 11 101 763 inhabitants at the beginning of 2071 which seems to be in harmony with all other projections. The share of people aged 65+ in 2071 predicted by B&K is lower than the same indicator predicted by CSU-2009 in 2066. The levels of net migration are assumed to be higher comparing to other projections, however the distribution of the net migration according to the different variants is pretty much the same as the one stated by CSU-2009. According to the B&K projection the total fertility rate will equal to 1.75 in 2070, and the life expectancy at birth will account for 84.9 for the males and 88.9 for the female part.

8.2 Data Collection

The main sources for the data collection were official authorities such as statistical databases of the major statistical offices. The projections of Eurostat and WPP together with all the variants and parameters were downloaded directly from the official databases of Eurostat and United Nations. However, there was a lack of data provided by the United Nations concerning parameters and age-groups.

The data of the CSU projections was partly collected from the official database. Another part covering particular parameters (total fertility rates, mortality rates, life expectancies) and variants of the projections was provided to the author directly form the workers of CZSO either by Terezie Štyglerová, or Michaela Němečková[1]. The data of the B&K projection with all the needed variants and parameters was provided directly to me by Boris Burcin. The data of B&K contained the 2009 projection with the three variants including all the age-groups and parameters except the crude death rate.

The projected data was evaluated against the data reported in the Eurostat's database containing the estimation of the Czech Republic's current population, which is in fact data collected by CZSO. This data corresponds to the data of all the projections regarding the base periods (except for WPP where the mid-period population sizes were used). However, it is important to notice that the data of real population or so-called observed values are just estimation, are not necessarily correct and may involve some inaccuracies as well

[1] The author of the thesis has passed the obligatory internship at CZSO during the second semester of the Master's study

Below you can see the table that summarizes the channels for obtaining the data of all the projections.

Table 12 Data sources of projections

Projection	Source
CSU – 2009	From T. Štyglerová and M. Němečková (Czech Statistical office) by e-mail, Czech Statistical office database https://www.czso.cz/csu/czso/databases-registers
CSU – 2013	From Štyglerová and Němečková (Czech Statistical office) by e-mail, Czech Statistical office database https://www.czso.cz/csu/czso/databases-registers
Eurostat – 2015	Eurostat Database https://ec.europa.eu/eurostat/data/database
WPP – 2015	United Nations Databese, World Population Prospects https://population.un.org/wpp/
B&K – 2009	From B. Burcin by e-mail

8.3 Data processing

All the work concerning the calculations and measuring was done in Microsoft Excel. Basically, it consists of two main parts:

1. The comparison of the individual projections between each other and with the real data observed for the period 2009-2017 and the visual representation of the comparison with the help of charts (all of them are presented in excel file) and constructing population pyramids to show the difference between real and projected values for the last available year (2017).
2. The measuring accuracy of the individual projections, variants and parameters with the help of the Keyfitz's "Quality of prediction index" (the standard one and with age groups), and the method of the Mean Absolute Percentage Error. After all calculations line and column charts were used to demonstrate the results (all of them are presented in excel file).

First of all, all the individual projections with their variants and parameters were collected. The projected data was taken up to the latest available year in the projection that to be able to present this data in graphs such that to observe the future development and compare the individual projections and their variants with the reality and between each other for the total number of population. Further, the data was presented separately for males and females, and divided into age categories in two ways. The first one was age groups from 0 to 14, 15 to 64, and 65+ taken for the total population and separately for males and females. The second one was age groups with the 10-years intervals from 0 to 100+ years taken for the total population and separately for males and females. Thus, this separation shows precisely in which age groups there might be the maximum or the minimum deviation from the real values. Beside this, the data for the following parameters was taken and presented in graphs: total fertility rate, net migration, life expectancy at birth by sex, and life expectancy at 65 by sex.

Also, population pyramids were constructed for the year 2017. Two overlapping population pyramids (based on the real data and on the predicted values) the best show the difference and the misfits of the real and predicted values for the last available year, also with the visual representation whether the projections are underestimated or overestimated, and to which ages those misfits belong. The population pyramids were constructed for the medium variants of each projection for the year 2017, thus in total there are 5 population pyramids. The real and projected population sizes were taken from the official databases as of January 1st except for United Nation projection, since its projection is presented as of July 1st. Hence, the real mid-period population size (2017-2018) was calculated to adequately compare WPP (United Nations) projection against the reality.

Keyfitz and MAPE were calculated for the years 2009-2017, since 2009 is the earliest available year for the given projections and 2017 is the last available year for the observed values from the official databases. They were computed for the all gender and age categories mentioned above, except for WPP projection which indexes of accuracy were found just for the total population size by sex because of the data unavailability. Since WPP constructs the population projection as of July 1 (the middle of the year), it is needed to find mid-period population sizes of observed values for each year to compare the observed values to the projected. In case of WPP there are just 2 mid-period population sizes that are available during 2015-2017, where one of them is the base size (2015-2016) and is the same as the observed, and the second (2016-2017) is the first projected period. Thus, this does not make much sense to analyze the projection just for the first projected year.

In addition, the Keyfitz's coefficient weighted by age groups was calculated for the following age categories: from 0 to 14, 15 to 64, and 65+ for the total, male and female population, and from 0 to 100+ years with 10-year step for the total, male and female population. As a result, the weighted index for the total population was found for each available year. In addition, simple Keyfitz was calculated separately for each age-specific fertility rate and age-specific life expectancy by gender for CSU-2013 and Eurostat projections.

Finally, the coefficients of Keyfitz and MAPE were presented in charts to show the deviation of projected values from the reality for each individual projection. The graphs were constructed for all the age groups and all the parameters available for the analysis. On the one hand, line charts were used to present the results of the Keyfitz's index since it shows the development of projection accuracy through the years. On the other hand, column charts were used to demonstrate the results of MAPE because it provides the average result of a particular projection for the whole period.

9 RESULTS

9.1 General findings

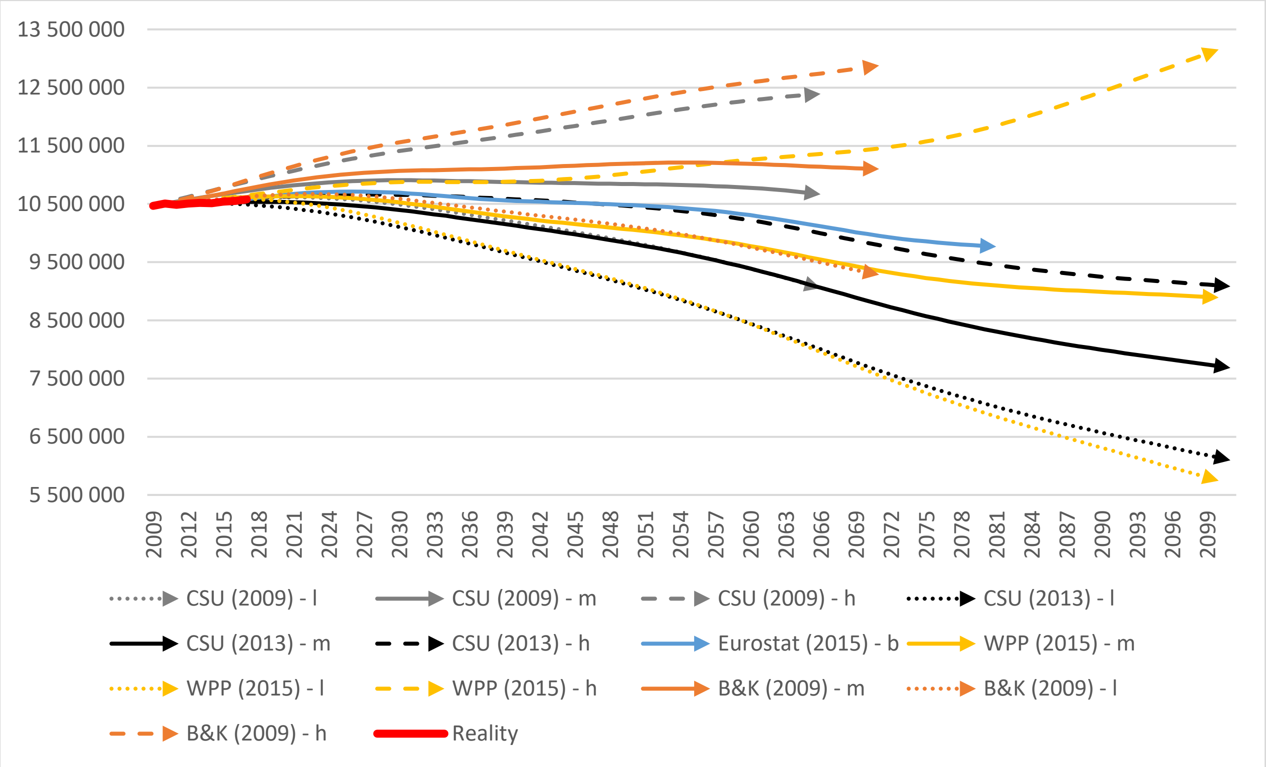
The general findings of the work are presented below and contain the overall trends that were detected in structure and development of the population projections.

9.1.1 Divergence of the projections

The figure 1 summarizes all the projections with their variants that were taken into the analysis at the initial stage. The figure reflects the probable future development of the Czech Republic population hold the current assumptions and trends of the population would stay the same. According to the different projections the population size would rise to 13 million people or would drop almost to 5.5 million by the year 2100.

With the time spent we can notice the divergence of the projections and their variant from the reality and within each other. This tendency can be followed in the figure 1. It is not surprising that the different projections are based on the different assumptions that lead to such varied results. However, the gap between the lowest variant and the highest variant out of all 13 is so huge up to the last available projected year, and, approximately, accounts for 7.5 million people, which is tremendously big number if we are talking about the population which for the current period consists of 10.5 million people. This fact may cause uncertainties and controversial opinions about the projections because if one of the variants would perform best of all in the future that would simultaneously mean that all other would be absolutely deviated from the reality.

Figure 1 Population size according to the individual projections vs. reality for the period 2009-2101

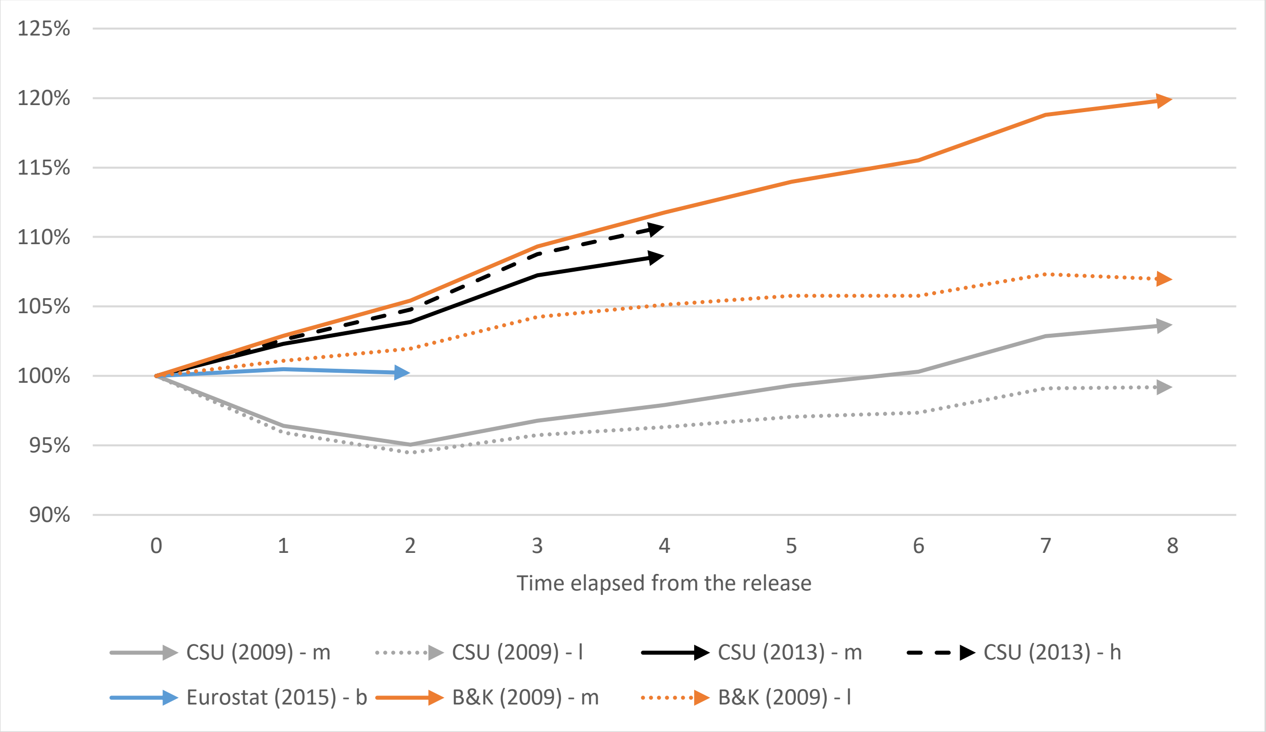


Source: CSU, Eurostat, UN, B&K, own elaboration

9.1.2 Deviation vs. time

Usually, the more years elapsed from the release of the projections, the larger the deviation from the real data is. The only exception might be Eurostat-2015 projection, which shows quite stable beginning of the projection (even though there are just 2 years available for the assessment). Compared to other projections' first 2 years Eurostat demonstrates the best results. For example, if we look at the Keyfitz's coefficient showing what is the deviation of the projected values from real data for the age group 90-99 years (figure 2), we can see that the deviation of almost each individual projection is getting higher with the time spent, while Eurostat projection is going around 100% line. The medium variant of B&K projection starts with the slight overestimation of 2.9 % after the first year from the release and finishes with 19.92 % of overestimation after the eighth year from the projection release, which makes the projections overestimated by 17 % more during the 7 years.

Figure 2 Keyfitz's index. Population aged 90-99 years according to the individual projections and the time elapsed from the projections' release



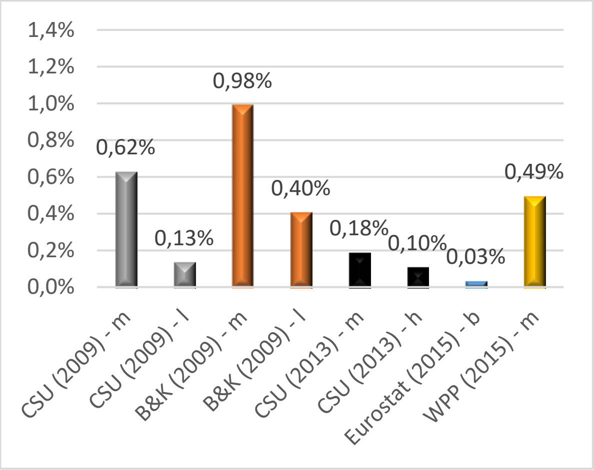
Source: CSU, Eurostat, B&K, own calculations

9.1.3 Male vs. female projections

Generally, the projections are better constructed for females; the accuracy of the male part is little bit lower. If we look at the figure 3 and figure 4, it can be clearly seen that MAPE is often higher for the male part of the projections for each individual projection. This tendency can be noticed throughout all the age groups and parameters.

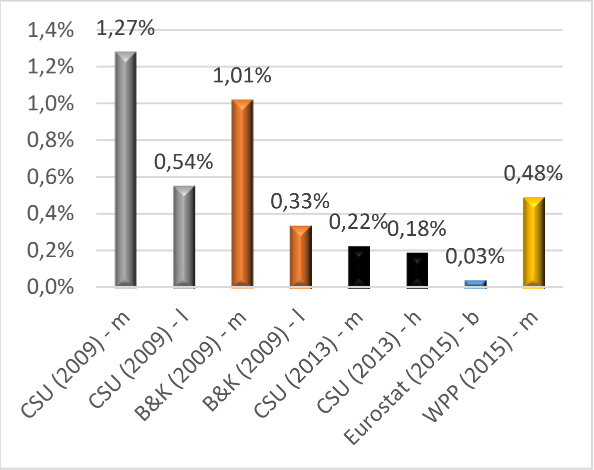
There are two things that stand behind such a tendency according to the author's opinion. On the one hand, we can take into account the sex ratio of the Czech Republic during the years 2009-2017 which fluctuates from 0.963 to 0.967 (Eurostat, 2019). The information that comes from this indicator is that there are more females than males in the population. Consequently, we can assume that it is harder to be more accurate for the male part of the population because of the higher error margin.

Figure 3 MAPE. Female population according to the individual projections for the whole period



Source: CSU, Eurostat, UN, B&K, own calculations

Figure 4 MAPE. Male population according to the individual projections for the whole period



Source: CSU, Eurostat, UN, B&K, own calculations

On the other hand, we can consider net migration by gender which was 15 512 for males and 9707 for females in 2016 (the last available year), (Eurostat, 2018). The number of men in net migration is considerably higher than the number of women, and since it is always hard to predict the future net migration, these two aspects result in higher deviation of the male part. Demographers predict male net migration with lower accuracy which consequently leads to lower accuracy of the whole projected male population.

9.1.4 Overestimation vs. underestimation

Generally, the variants of the projections (even low variants) are overestimated in comparison to reality. The Keyfitz's coefficient tells us the information about whether the projection is overestimated or underestimated. The table 13 demonstrates the results of the Keyfitz's coefficient for the total population by individual projections and years. We can see that almost all coefficients are higher than 100% which means that the projected value is overestimated (without color), and just 7 coefficients out of 51 are lower than 100% which means that the projected value is underestimated (light yellow color). Also, from this table we can learn that the Keyfiz's index of Eurostat performance equals to 100.00% which indicates the excellent performance and very low deviation from reality (light green color).

The same will be true if we look at male and female population separately. The evidence is that most of the variants are overestimated, and most likely that this overestimation is belonged to older projections like CSU-2009 and B&K-2009. This can be explained by positive demographic developments before the year 2009, and, indeed, if we look at the evolution of the Czech Republic's population over the years, we will see that there was a sharp increase of the population size after the year 2003 continuing till 2011 which provoke demographers to construct the projections based on higher assumptions. After the year 2011 the population size increases just slightly, what resulted in newer projections (CSU-2013, Eurostat -2015) assumptions, hence lower overestimation or even underestimation.

Table 13 Keyfitz's index. Total population according to the individual projections

Projection	2009	2010	2011	2012	2013	2014	2015	2016	2017
CSU (2009) - m	**100,00%**	99,97%	100,49%	100,64%	100,85%	101,19%	101,23%	101,36%	101,38%
CSU (2009) - l	**100,00%**	99,85%	100,24%	100,25%	100,32%	100,51%	100,39%	100,35%	100,20%
CSU (2013) - m					**100,00%**	100,11%	99,92%	99,81%	99,59%
CSU (2013) - h					**100,00%**	100,22%	100,14%	100,15%	100,05%
Eurostat (2015) - b							**100,00%**	100,00%	100,00%
B&K (2009) - m	**100,00%**	99,92%	100,41%	100,55%	100,80%	101,20%	101,35%	101,59%	101,73%
B&K (2009) - l	**100,00%**	99,81%	100,19%	100,21%	100,31%	100,54%	100,49%	100,51%	100,43%
WPP (2015) - m							100,55%	100,42%	

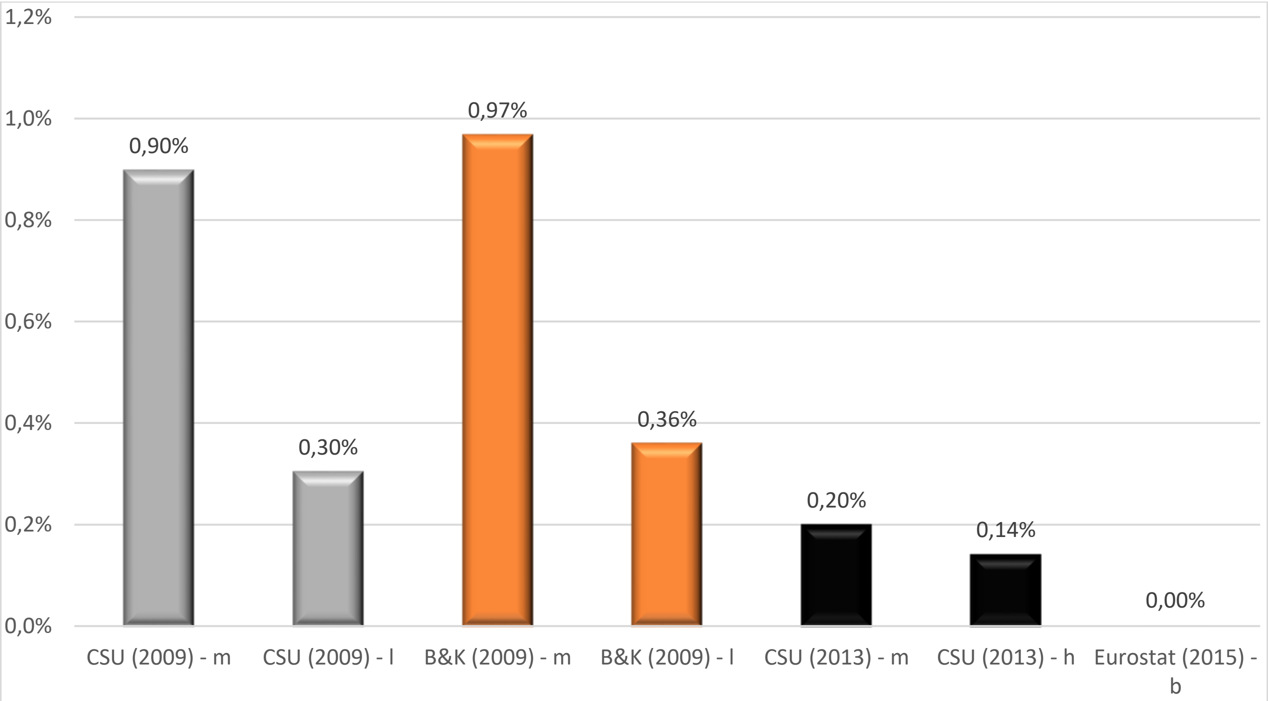
Source: CSU, Eurostat, UN, B&K, own calculations

9.2 Individual projections

It makes sense to start with the comparison of the individual projections with their variants to be able to see how different they are. The figure 5 demonstrates MAPE of the total population according to the projections for the whole period. From this figure we can learn that Eurostat-2015 has absolutely the very good performance with the average error of 0.001%. However, this result cannot be compared to the results of other projections (MAPE is calculated just for the period of 2 years) as well as the result of CSU-2013 (MAPE is calculated for the period of 4 years) since the projections were published during the different calendar years. The projection of Eurostat is one of the newest and it is possible to evaluate it just for the period 2015-2017. The results of such evaluation will be discussed in the sub-chapter 9.5.

The projections that were released in the same year can be compared from this chart. We can see that the average error of the projection B&K and CSU-2009 look pretty much the same with the little bit more accurate prediction by CSU. The medium variants of CSU-2009 and B&K-2009 are turned to be the least accurate among other projections for the whole period. The low variants of CSU-2009 and B&K-2009 both look better since, as was highlighted above, the overestimation was typical for the old projections, where the low variants surely win over the medium ones.

Figure 5 MAPE. Total population according to the individual projections for the whole period



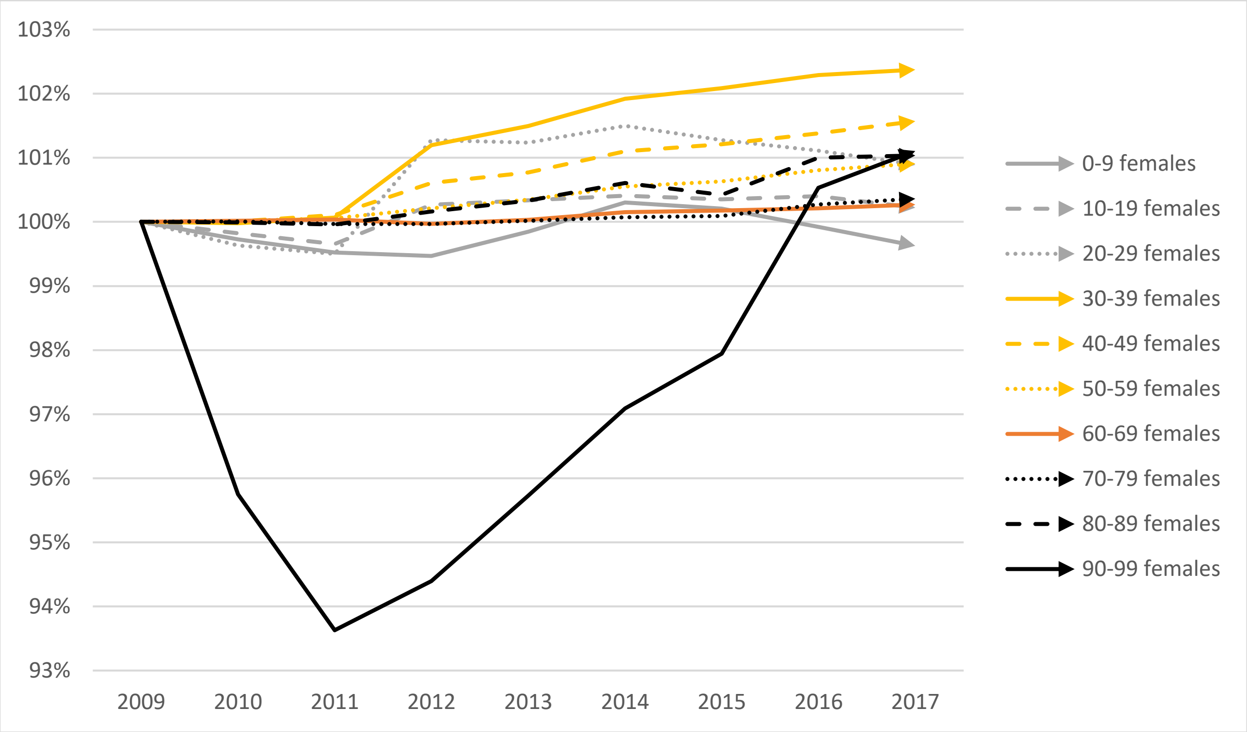
Source: CSU, Eurostat, UN, B&K, own calculations

9.2.1 CSU 2009

CSU 2009 is better constructed for females. On average, the low variant performs better for the whole population (figure 5), but it cannot be surely said whether the medium or low variant is better. For example, the medium variant shows very accurate result for the age category 0-9 (the Keyfitz's index fluctuates very close to 100% throughout the whole evaluation period) years both for males and females; the low variant is underestimated for this age category. However, for the age categories 20-29 and 40-49 years of females, the medium variant proves worse results than the low variant.

CSU-2009 shows the best performance for the age categories 10-19 and 50-79 years both sexes (see figure 6 and figure 7). Also, it is good for the females aged 80-89 (see figure 6). The projection is bad produced for the age categories from 20 to 49 years (see figure 6 and 7) both sexes and 80-89 years of the male part of the population. As for parameters, CSU-2009 follows the general trait of all the projections, but the low variant demonstrates better results for total fertility rate and crude death rate. From the figure 7 we can observe that the Keyfitz's coefficient of female population by age groups belongs to the interval [99.47%; 102.38%] (age group 90-99 years is not included due to significantly higher deviation). Figure 8 shows that the deviation interval of male population equals to [99.8%; 104.47%]. According to the graphs mentioned above the both variants are mostly overestimated comparing to the reality.

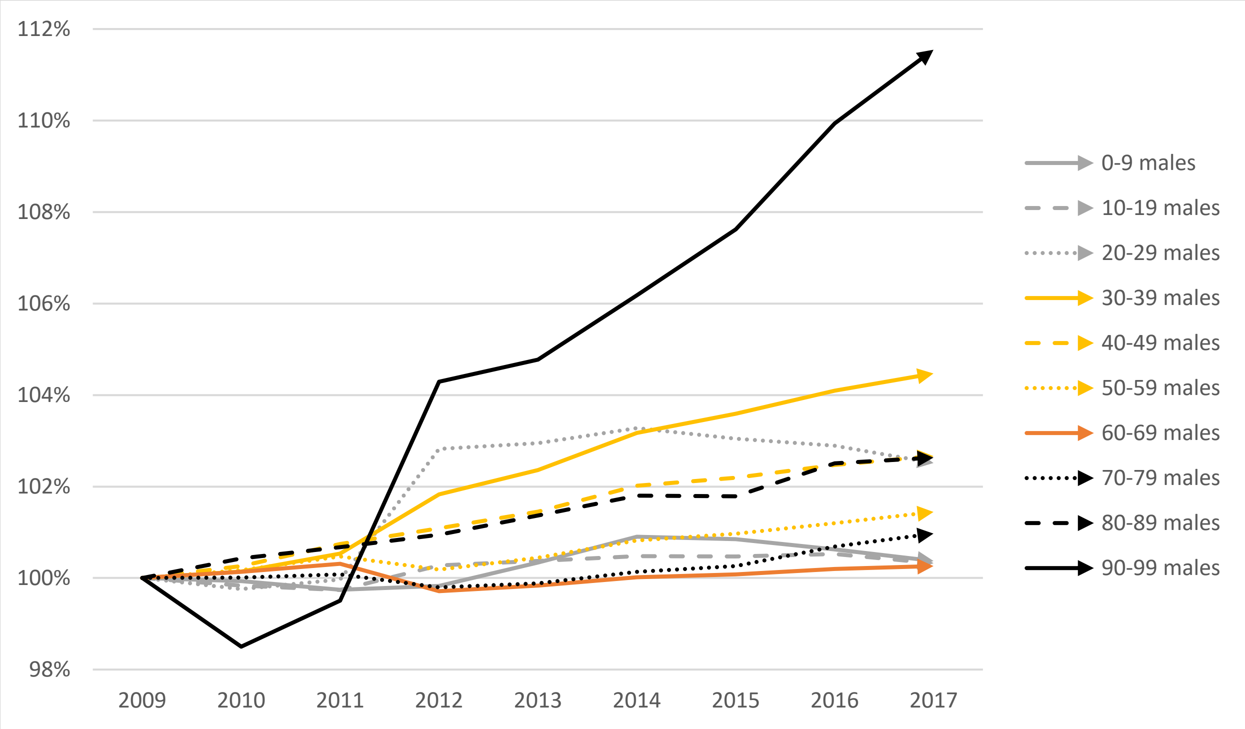
Figure 6 Keyfitz's index. Female population based on age groups, CSU (2009) – medium variant



Source: own calculations, CSU

The figure 8 presents the population pyramid in 2017 including the real values and the projected values of CSU-2009 – medium variant. It is visible that in 2017 the projection is mainly overestimated despite the ages 0-1 and 6-10. The upper part of the pyramid more matches the reality than the bottom and middle part. The female part is constructed more correctly; male part is much overestimated at the age from 28 to 40 years.

Figure 7 Keyfitz's index. Male population based on age groups, CSU (2009) - medium variant



Source: own calculations, CSU

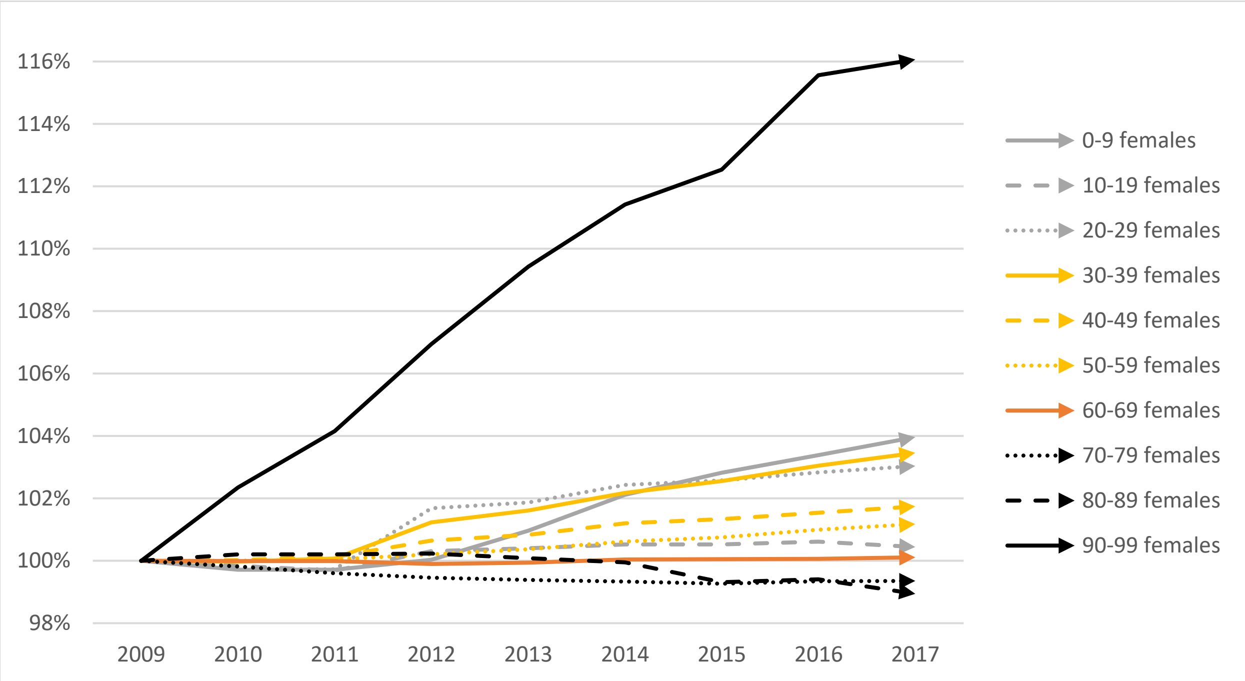
Figure 8 Real population structure vs. projected population size by CSU (2009) - medium variant, 2017

Source: own elaboration, CSU

9.2.2 B&K 2009

B&K-2009 seems to be the worst projection among others. On overall, the low variant of B&K-2009 projection shows better result than the medium variant (see figure 5). However, the Keyfitz's coefficient by age category demonstrates bad results for males, females and total population both for medium and low variant. The exception is the low variant for males involving the age categories 0-14; 15-64; 65+ years.

Figure 9 Keyfitz's index. Female population based on age groups, B&K (2009) - medium variant



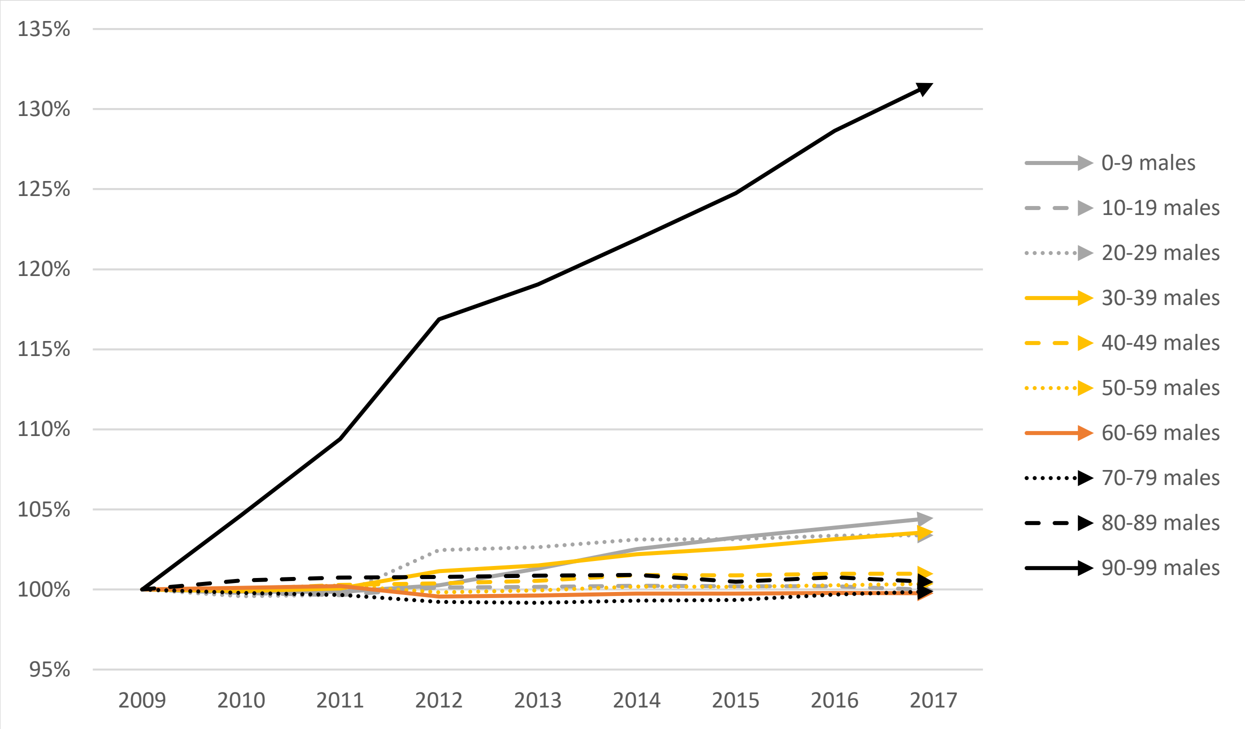
Source: own calculations, B&K

The projection is bad constructed for the age categories 0-9, 20-49, 90-99 years for females (see figure 9). According to the figure 10 the situation looks the same for the male part except the age group 40-49 years. The least accurate age groups for the males are 0-9 and 20-39. The deviation for the age group 0-9 years rises up to 4.47 %, 20-29 years to 3.38%, 30-39 years to 3.61% in 2017. Actually, the similar tendency is observed for all the projections, where the least accurate age groups belong to the same age intervals.

The low variant shows not very good results also for age categories 60-69 males, 70-79 males and females, 80-89 males and females. The medium variant has worse results for the age category 100+ years. As for parameters, B&K-2009 demonstrates the worst result for total fertility rate. Life expectancy at birth is underestimated by low variant both for males and females.

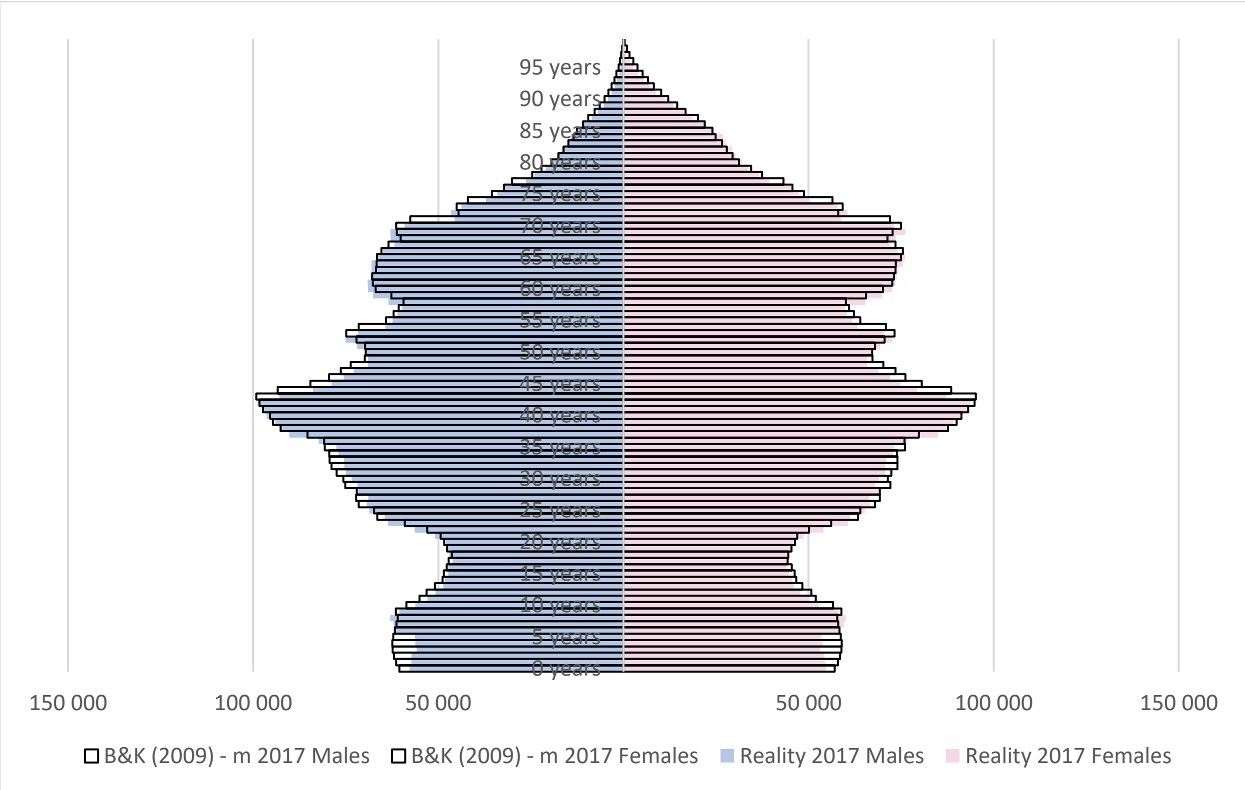
From the figure 9 we can observe that the Keyfitz's coefficient of female population by age groups belongs to the interval [98.94%; 103.96%] (age group 90-99 years is not included due to significantly higher deviation). The deviation interval of the male part (figure 10) equals to [99.89%; 104.47%]. Those two intervals are the widest ones comparing to other projections' maximum deviation.

Figure 10 Keyfitz's index. Male population based on age groups, B&K (2009) - medium variant



Source: own calculations, B&K

Figure 11 Real population structure vs. projected population size by B&K (2009) - medium variant, 2017



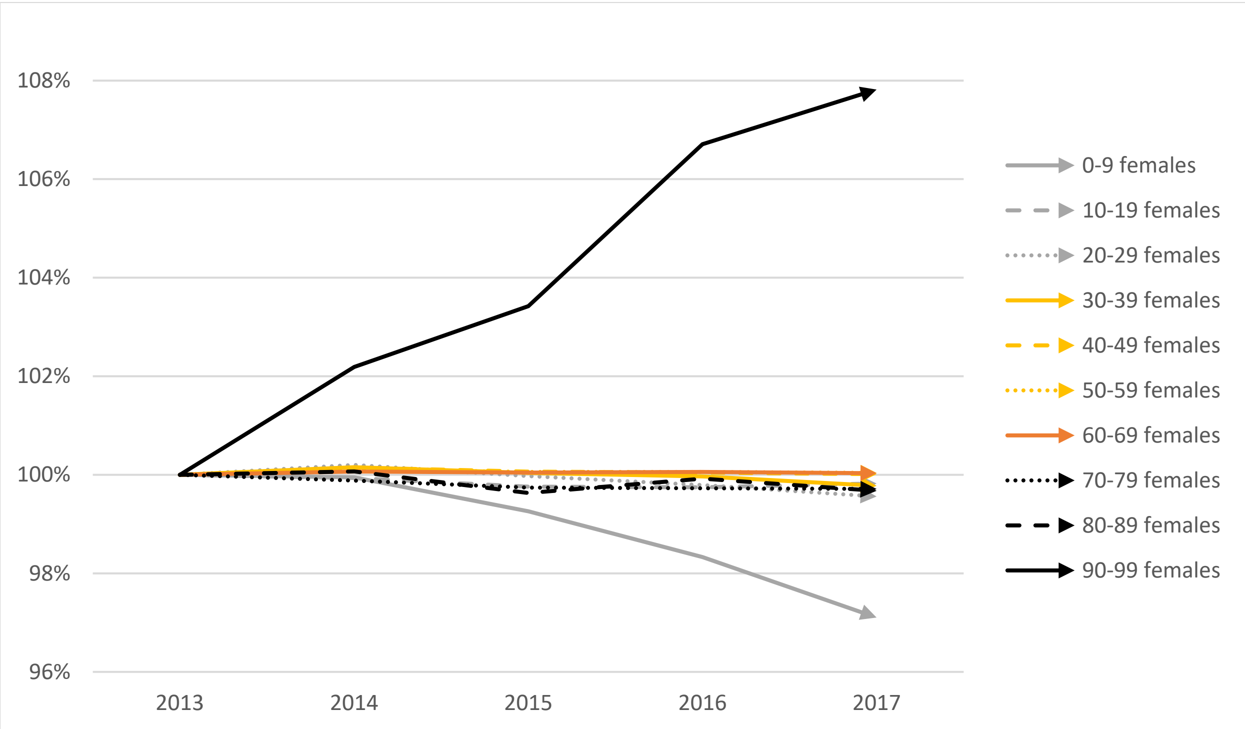
Source: own elaboration, B&K

On the figure 11 we can see the population pyramid constructed for the B&K-2009 – medium variant projection in 2017. This pyramid is the worst among others since the mismatching is visible so much throughout the all ages both on the male and female part. It seems like the projection is constructed with 1-year overtake, which means that what was projected for the age 40 must be shifted to the age 39 and so on. That would make the projection to match the reality better. The most inaccurate part is male and female population at the age of 0-5 years. There is a huge overestimation of the number of births with followed up by the huge overestimation in young age groups.

9.2.3 CSU 2013

CSU-2013 is well constructed both for males and females, and it is proved by the medium variant as well as high variant. The figure 5 presents MAPE for the whole period comparing medium and high variant of the projection, where we see that the difference between them is not that high. CSU-2013 shows quite stable results throughout all the age categories. The exceptions are the age category 0-9 years both sexes (figure 12 and 13) and both variants, males aged 80-89 years old (figure 13) both variants, and 90-99 years both sexes (that can be visible on the figure 12 and 13) and both variants. In particular, from the figure 13 we can see the pure example of the difference of older age groups comparing to other age groups. It proves that with elder ages the deviation is getting higher, especially for the age group from 90 to 99 years where the projection is overestimated by 11.21% comparing to the reality.

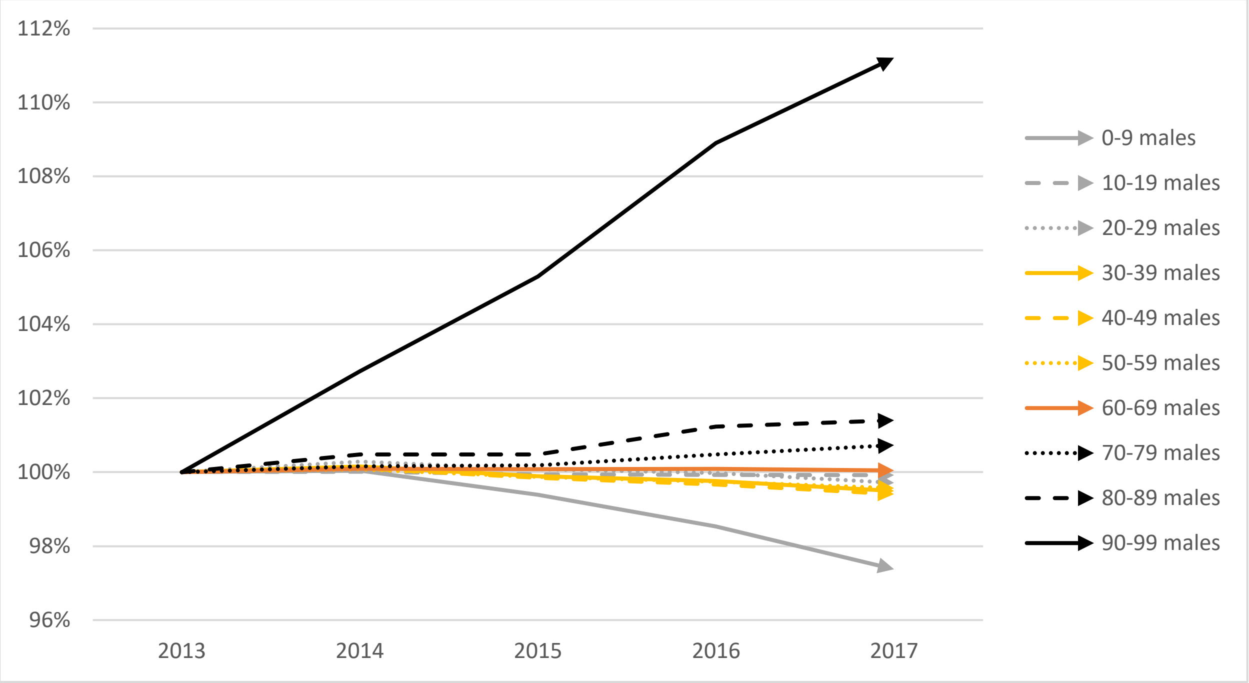
Figure 12 Keyfitz's index. Female population based on age groups, CSU (2013) - medium variant



Source: own calculations, CSU

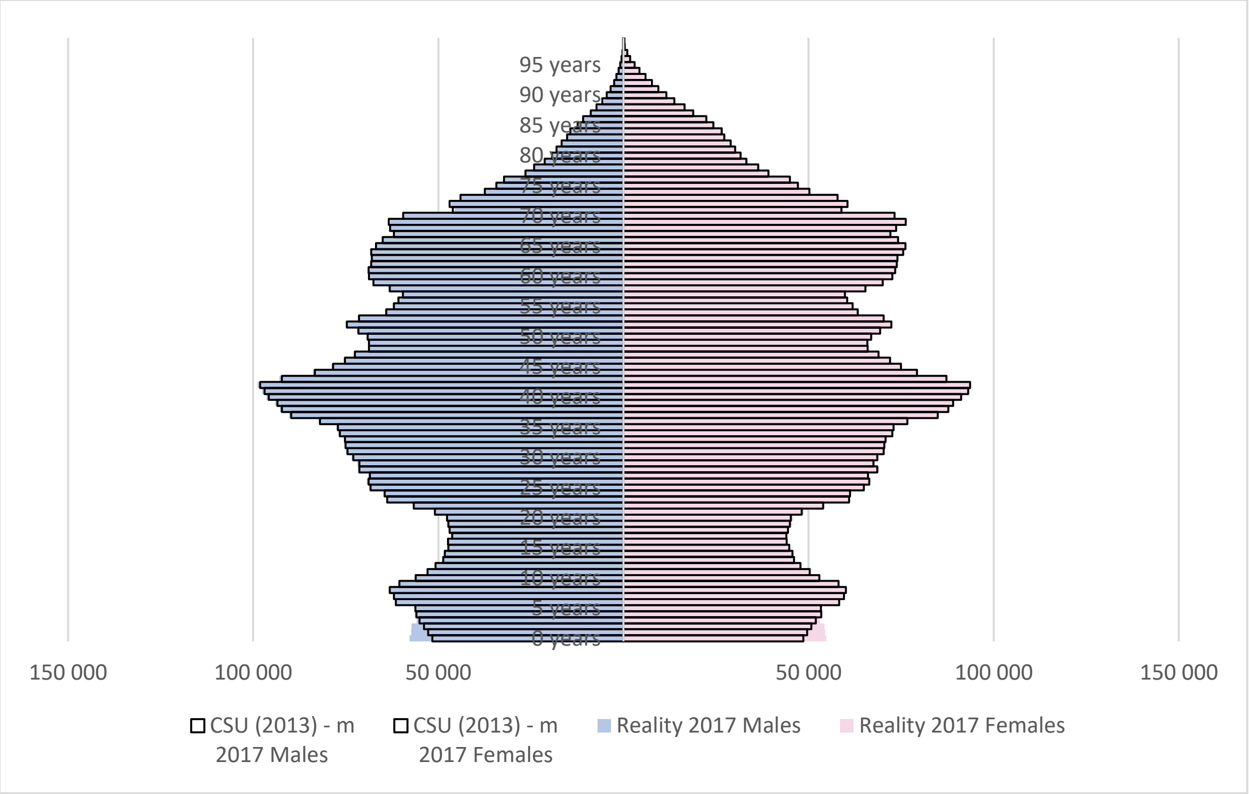
From the figure 12 and 13 we can observe that the Keyfitz's coefficient of female population by age groups for the whole period belongs to the interval [97.10%; 100.03%] (age group 90-99 years is not included due to significantly higher deviation). Analogically for the male part the deviation interval equals to [97.38%; 101.40%] what is a really good indicator of performance. Also, as evidence to the fact mentioned in the chapter 9.1.4 both variants of CSU-2013 go above and below the real numbers, they are not that much overestimated. On the contrary, the female part of the population is mostly underestimated comparing to the reality.

Figure 13 Keyfitz's index. Male population based on age groups, CSU (2013) - medium variant



Source: own calculations, CSU

Figure 14 Real population structure vs. projected population size by CSU (2013) - medium variant, 2017



Source: own elaboration, CSU

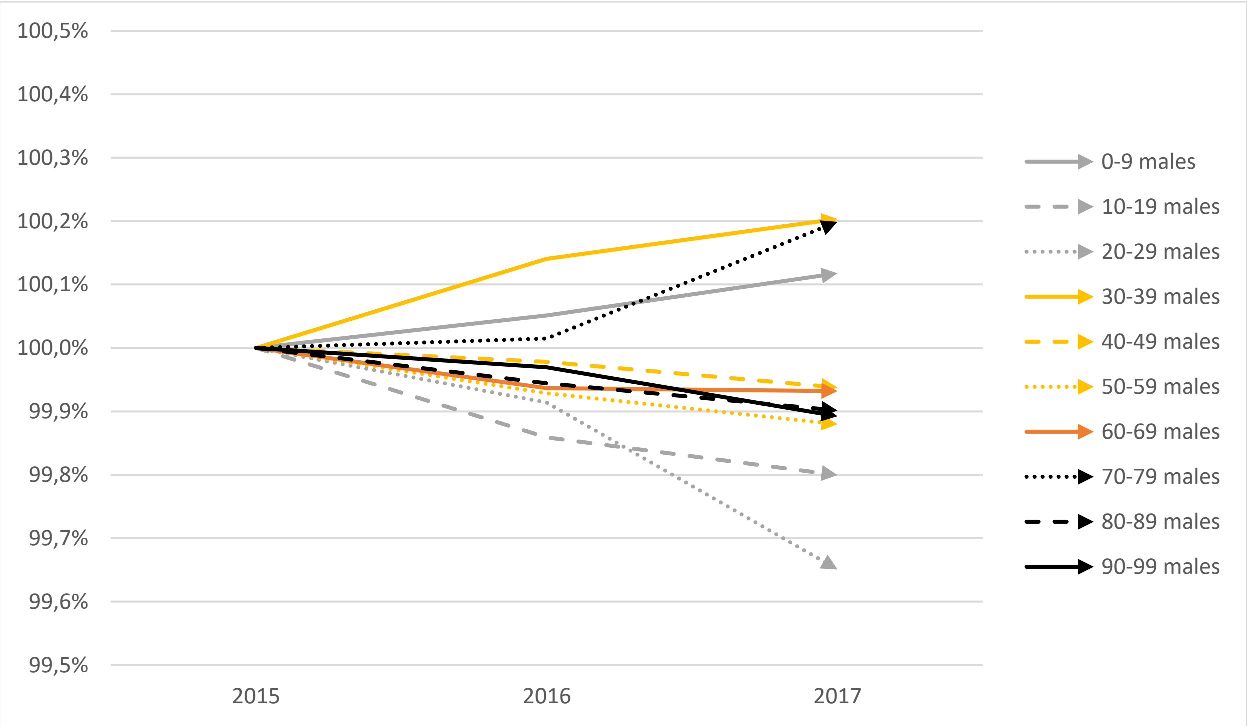
As for parameters, both variants of CSU-2013 follows the general trait of all the projections, however, total fertility rate is underestimated by both variants, and it is getting worse with the time elapsed from the projection release. Also, the high variant provides the least accurate result for the net migration among other projections with the average error of 392.5% for the whole period.

The figure 14 shows the population pyramid of CSU-2013 – medium variant in 2013. On the bottom of the pyramid it is visible that the projection is underestimated up to the age 3, which absolutely corresponds to the higher deviation of 0-9 age group in 2017.

9.2.4 Eurostat 2015

Eurostat has the most stable and the most accurate projection so far. Despite the fact that it is one of the newest and it is possible to evaluate it just for the period 2015-2017, it is still the best in comparison with the first 2 years of other projections. If we look at the graphs with the Keyfitz's coefficient by different age categories according to the individual projections and the time elapsed from the projections release, we will notice that Eurostat 2015 has the lowest deviation from the reality for the first 2 years spent (discussed in the subchapter 9.5). Eurostat 2015 demonstrates excellent coefficients throughout the calculations both for Keyfitz and MAPE (just for 2 years). The only exception is life expectancy at the age 65 for males and females. However, Eurostat looks much better even for the age categories 90-99 and 100+ years old.

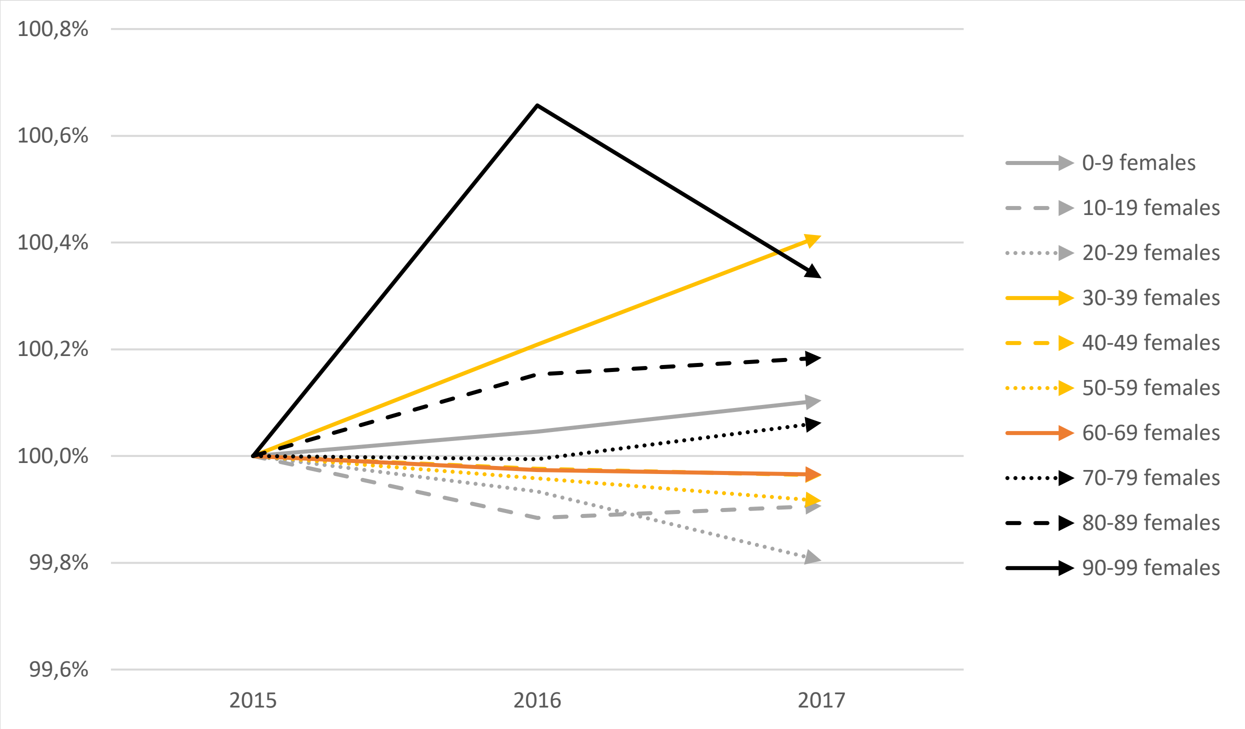
Figure 15 Keyfitz's index. Male population based on age groups, Eurostat (2015) - baseline variant



Source: own calculations, Eurostat

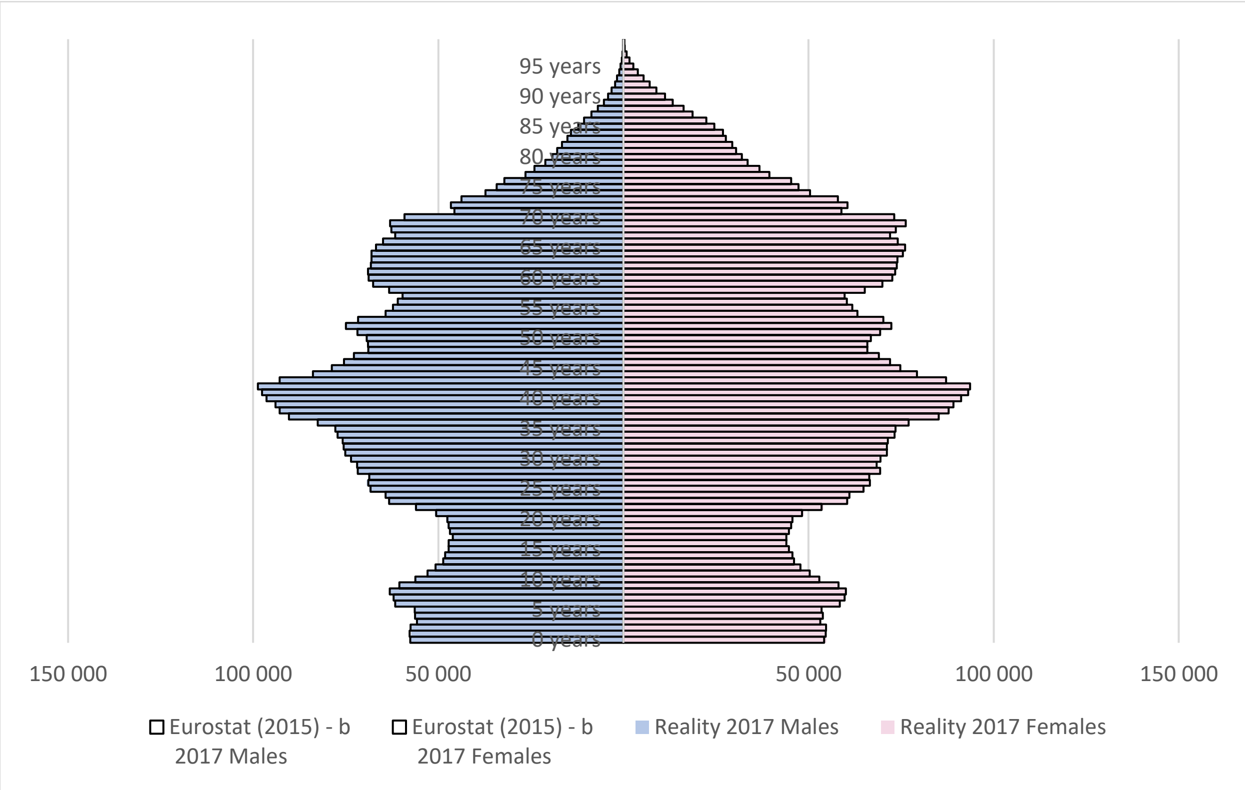
The figure 15 represents the Keyfitz's coefficient of male population by age group of Eurostat projection for the period 2015-2017. We can see that the most problematic parts are 20-29 and 30-39 age groups. Even though they are problematic, the Keyfitz's index belongs to the interval of [99.65%; 100.20%] which means that the deviation does not exceed 0.35% for the lower boundary and 0.2% for the upper boundary.

Figure 16 Keyfitz's index. Female population based on age groups, Eurostat (2015) - baseline variant



Source: own calculations, Eurostat

Figure 17 Real population structure vs. projected population size by Eurostat (2015) - baseline projection, 2017



Source: own elaboration, Eurostat

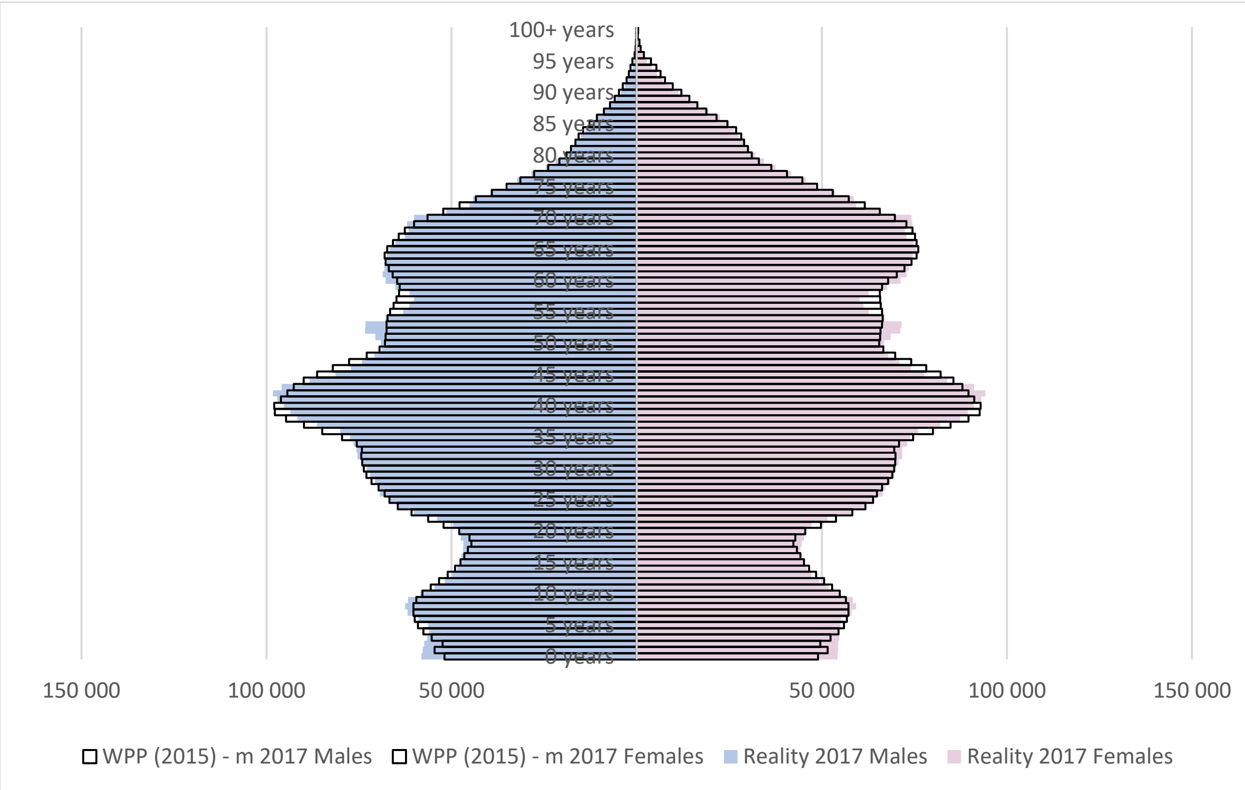
The figure 16 demonstrates the Keyfitz's index for the female part of the population. The least accurate age groups are 30-39 and 90-99 years with the deviation of 0.41% and 0.33% respectively for the last available year. The interval boundaries equal to [99.80%;100.66%] saying to us that the deviation does not exceed 0.2% for the lower boundary and 0.66% for the upper boundary, in fact which is lower for the last available year and accounts for 0.41%. That is a very good indicator of performance.

The figure 17 represents the population pyramid of Eurostat-2015 – medium variant vs. reality in 2017. It clearly seen that the projection demonstrates a perfect result comparing to the real numbers. It is hardly noticeable any mismatches in the construction of the projection by gender and age, which is absolutely corresponding to the results of Keyfitz and MAPE.

9.2.5 WPP 2015

Not that much work was done with the United Nations projection because of the data availability. The evaluation was carried out just for the total population, male population and female population, since it was possible to assess just the first year after the projection release which does not make much sense (United Nations prepare projections as of July 1st, it is needed to calculate the mid-period population sizes that to be able to compare the WPP projection to the reality). WPP-2015 demonstrates the average results among all other projections. According to Keyfitz and MAPE, the projection is overestimated for the first year for the total population, as well as for males and females separately.

Figure 18 Real population structure vs. projected population size by WPP (2015) - medium variant, 2017



Source: own elaboration, United Nations

On the figure 18 we can see that different part of the projections is either overestimated or underestimated. The upper part of the pyramid better matches the reality. On average, female part is by little better constructed. As it is visible at the graph the most mismatching part is population at the ages of 48-59 years old and the youngest ages approximately from 0 to 4 years. It is important to stress out that the real data was calculated as a mid-period population size between 2017 and 2018 to be able somehow to compare the projection that is produced as of 1st July. So, the output that is presented in graph might be biased and might be less or more accurate in fact.

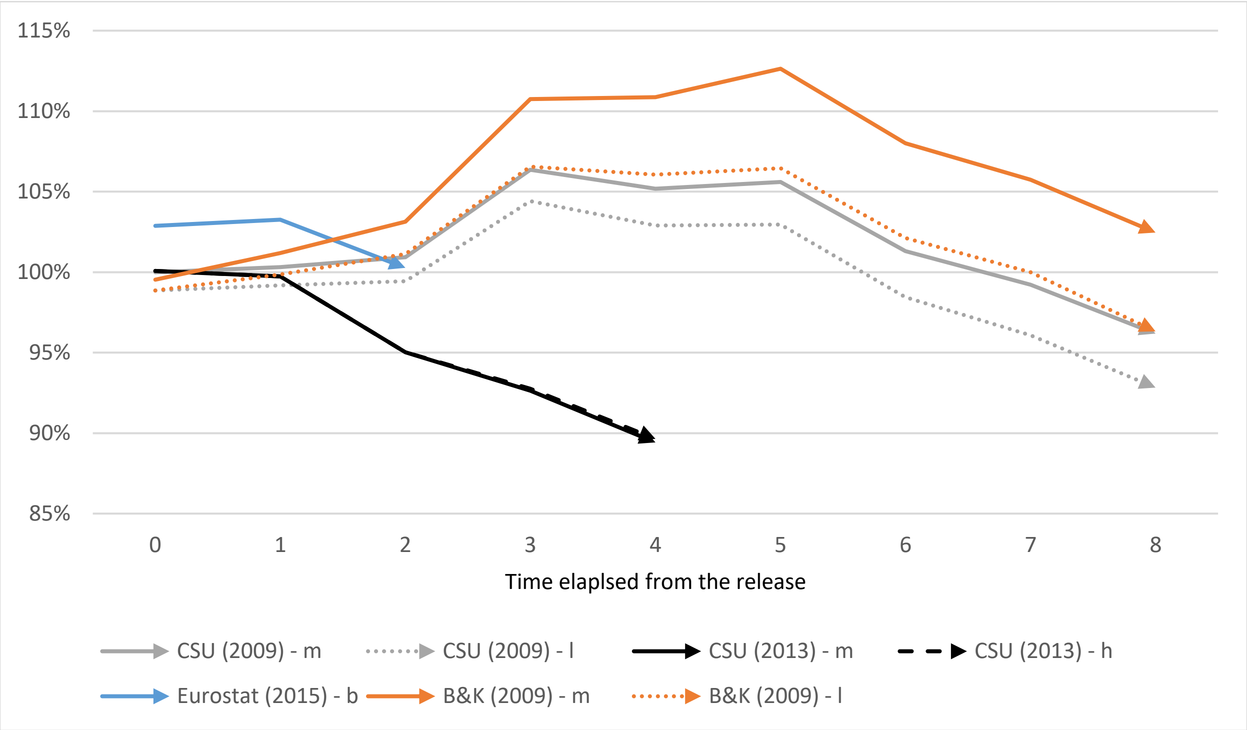
9.3 Parameters

9.3.1 Total fertility rate

Regarding the parameters of the projections, there are almost no problems with total fertility rate and life expectancy at birth. The figure 19 represents the Keyfitz's coefficient of total fertility rate by the individual projections and the years spent from the projections release. The index belongs to the interval of [89.42%; 112.64%] for the whole period. All the producers have different initial total fertility rate as the base for the projection. The highest one belongs to Eurostat projection, the low variant of B&K-2009 and the low variant of CSU-2009 start at the lowest level.

It is noticeable that the projections of CSU-2013 both variants have significantly underestimated total fertility rate comparing to other projections and to the reality, especially with more time spent after the projections' release. Moreover, CSU produced identical projection no matter of low or medium variant. B&K-2009 medium variant projection considerably overestimates total fertility rate with the deviation more than 10% from the observed data. We can see that up to the second year after the projections' release the projected values are fluctuated around the real values, however, after the second year after the release they rise and go down not in accordance with the reality.

Figure 19 Keyfitzs's index. Total fertility rate according to the individual projections and the time elapsed from the projections release



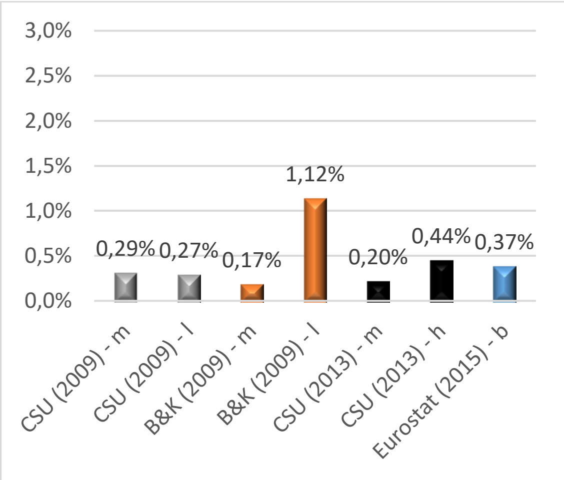
Source: CSU, Eurostat, B&K, own calculations

Also, from the figure 19 we can see that all the projections have the downcast shape of the curve, for the older projections this trend is observed after the 5th year elapsed from the release, and for the newer projections it is visible at once. Actually, after the year 2013 the total fertility rate in the Czech Republic rose up sharply from 1.45 to 1.63 in 2017. By this fact the downcast shape of the curves can be explained, since almost all the projections predicted total fertility rate below the value 1.63. B&K-2009 medium variant was the closest projection to the real value in 2017 (1.67), and the Eurostat projection was absolutely equal to the real total fertility rate in the Czech Republic in 2017.

9.3.2 Life expectancy

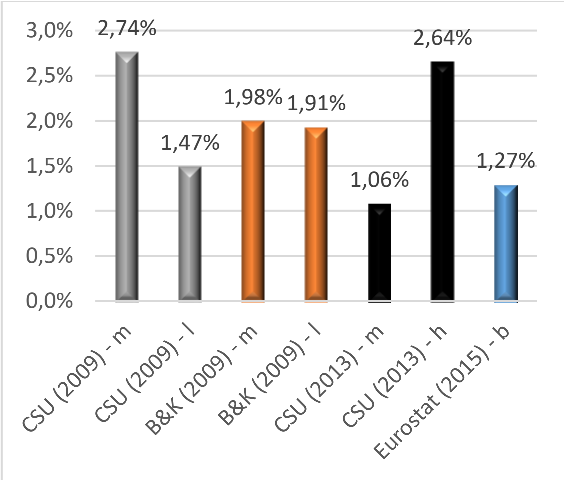
The error of the projected life expectancy at birth is low for all the projections. Unlike other age groups and parameters, the medium variant of B&K-2009 performs best of all for the male life expectancy at birth with the average deviation 0.17% for the whole period according to MAPE, the worst one is also B&K-2009 but with the low variant and the deviation of 1.12%.

Figure 20 MAPE. Male life expectancy at birth according to the individual projections for the whole period



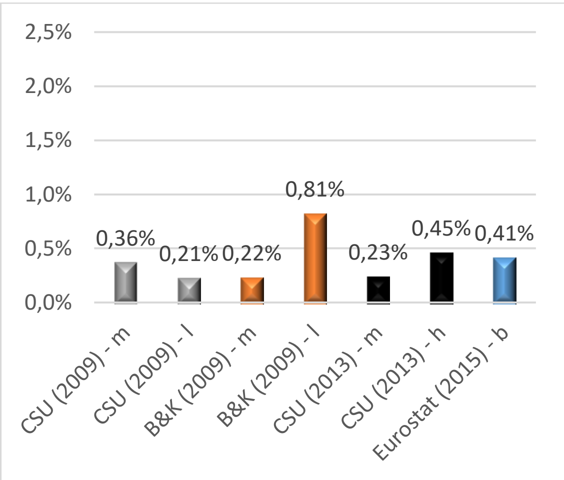
Source: CSU, Eurostat, B&K, own calculations

Figure 21 MAPE. Male life expectancy at 65 according to the individual projections for the whole period



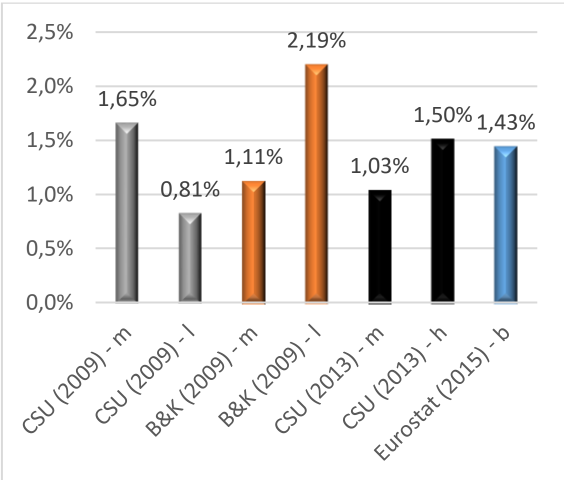
Source: CSU, Eurostat, B&K, own calculations

Figure 22 MAPE. Female life expectancy at birth according to the individual projections for the whole period



Source: CSU, Eurostat, B&K, own calculations

Figure 23 MAPE. Female life expectancy at 65 according to the individual projections for the whole period



Source: CSU, Eurostat, UN, own calculations

As for the female part, we can see that all the projections, more or less, show good results for the life expectancy at birth (figure 22), the low variant of CSU-2009, the medium variant of CSU-2013 and the medium variant of B&K-2009 perform best of all with the deviation of 0.21%, 0.23% and 0.22% respectively for the whole period. Interestingly, this parameter is predicted more accurate for the males by several projections (CSU-2013 medium variant, B&K-2009 medium variant). B&K-2009 low variant demonstrates the highest deviation of 0.81%, which is, in fact, better than the worst result of the male life expectancy at birth.

The accuracy of life expectancy at the age 65 is low both for the males and the females with a little bit better results for the female part. The figure 20 and the figure 21 demonstrate MAPE of male life expectancy at birth and male life expectancy at 65 for the whole period respectively. It is clearly seen that the deviation of life expectancy at 65 is considerably higher for all the projection and variants. The same is true for the female part of the population (figure 22 and figure 23). There are two reasons that stand behind the lower accuracy of the life expectancy at 65. The first one is that this parameter deals with the high error margin (we need to be more precise because it is less and less years left to live up to death). The second is about changing mortality rates, especially for higher ages that is hard to predict precise in advance.

CSU-2013 – medium variant shows the best results among others for the male life expectancy at 65 with the deviation 1.06%, CSU-2009 – medium variant has the highest deviation of 2.74%. CSU-2009 low variant shows the best result among others for the female life expectancy at 65 with the deviation accounted for 0.81% (better than males). The least accurate projected life expectancy at 65 for females was produced by the low variant of B&K-2009 projection with the deviation of 2.19% for the whole period (better than males). It is interesting to notice that the average deviation of the Eurostat projection is relatively high taken into account the fact that the average error was evaluated just for the period of 2 years.

9.3.3 Net migration

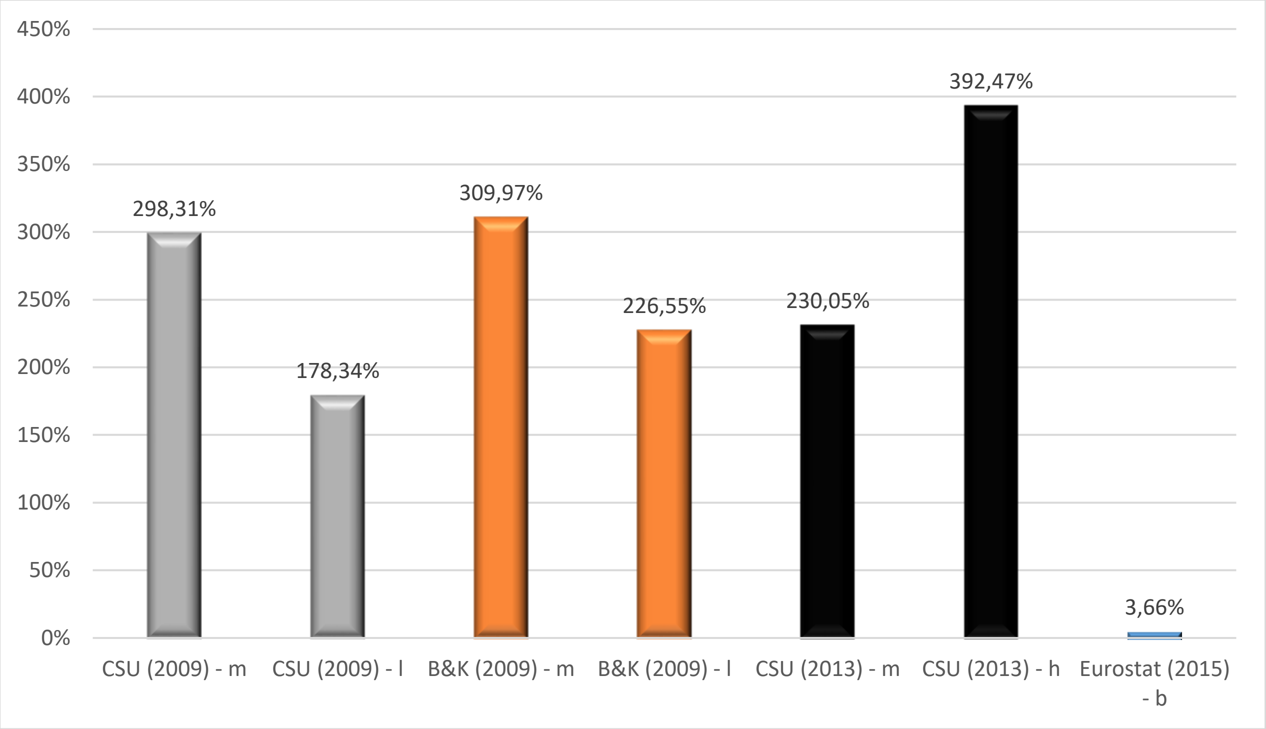
The most unpredictable and problematic parameter of all the projections is net migration. It is not surprising that the deviation of this parameter can reach very high numbers, since there are so many reasons that can stand behind it. Net migration is the flow that cannot be logically explained because it involves human factors, like motivation to move and necessity to move, are the trends that not fully predictable in advance.

The figure 24 represents the average deviation of projected values from real values for the whole period. The chart shows that all the projections proved to be poorly constructed for the parameter of net migration. In case of CSU-2013 high variant the error even almost reaches the deviation of 400%. Eurostat performs very well for the period of 2 years with the outstanding result of 3.6%.

It makes sense to present the results of the Keyfitz's index as well (figure 25), since MAPE shows juts the average error for the give period, and it is known that outliers (which is very typical for net migration) can distort the results if we speak about averages.

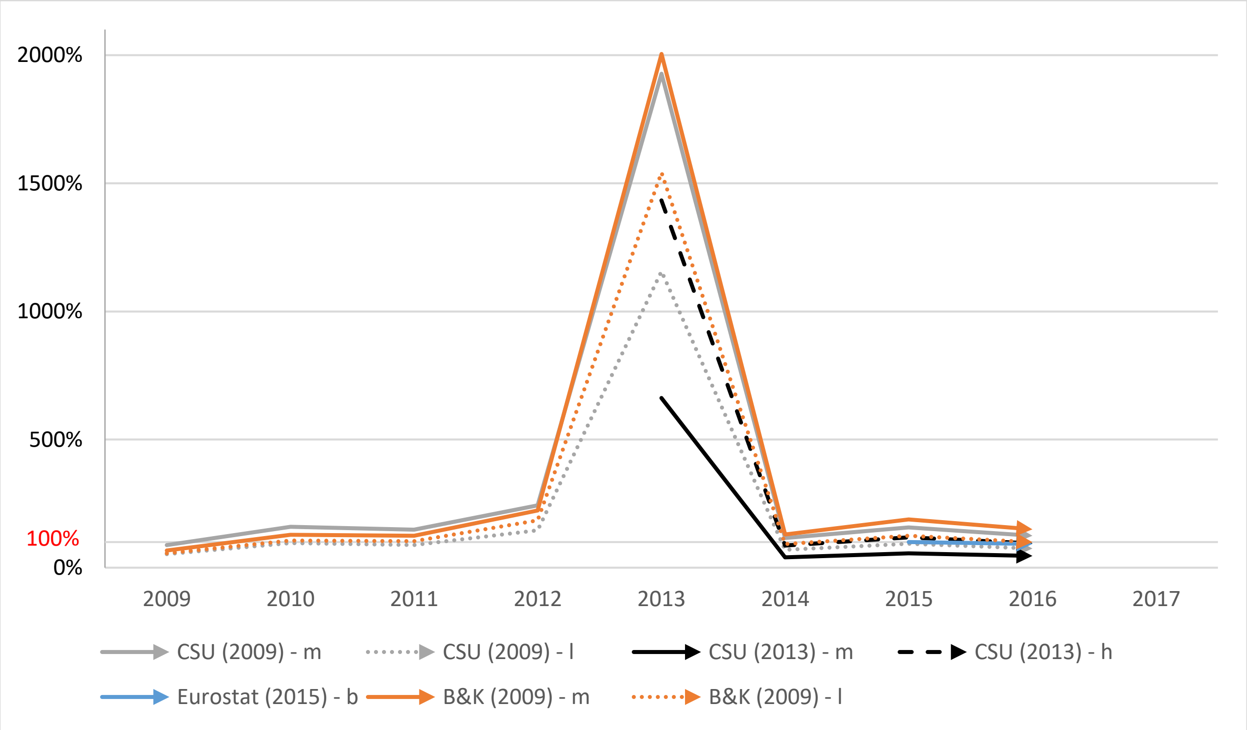
It also makes sense to construct the below chart with the horizontal axis equaled to the concrete calendar year (2009-2017) unlike all the previous charts where time elapsed from the projections' release was used as the horizontal axis. Here, interesting situation is observed providing the extremely sharp increase in the deviation reaching even 2000% for some projections (B&K-2009 medium variant, CSU-2009 medium variant) in 2013. In fact, this is the year that absolutely destroyed MAPE outcomes for all the projections, except Eurostat. Now it is clear why Eurostat performs such a good result, since its projection was released after the critical year in 2015.

Figure 24 MAPE. Net migration according to the individual projections for the whole period



Source: CSU, Eurostat, B&K, own calculations

Figure 25 Keyfitz's index. Net migration according to the individual projections and the years



Source: CSU, Eurostat, B&K, own calculations

The reason that stands behind such a huge gap is the negative balance of net migration that was observed in the Czech Republic in 2013 and equaled to -1297 (table 14). According to International Migration Outlook (2015), "*for the first time since 2001, the Czech Republic experienced a net migration outflow in 2013, the migration outflow of 31000 persons exceeded the migration inflow by more than 1000 persons*". In 2013 there were high number of emigrants (comparing to previous years) together with lower than usual number of immigrants, what causes the negative balance of net migration, and surely was not able to predict. Hence, the low variants have the advantage against the medium and high variant within the projection to demonstrate better accuracy; however, the two worst results belong not to the high variants but to the medium variants of B&K-2009 and CSU-2009.

Table 14 Net migration according to the individual projections and reality by years

Net migration	2009	2010	2011	2012	2013	2014	2015	2016	2017
CSU (2009) - m	25 000	25 000	25 000	25 000	25 000	25 000	25 000	25 000	25 000
CSU (2009) - l	15 000	15 000	15 000	15 000	15 000	15 000	15 000	15 000	15 000
CSU (2013) - m					8 587	8 743	8 934	9 150	9 378
CSU (2013) - h					18 587	18 707	18 864	19 045	19 239
Eurostat (2015) - b							15 983	18 601	23 741
B&K (2009) - m	19 000	20 000	21 000	23 000	26 000	28 000	30 000	30 000	30 000
B&K (2009) - l	16 000	16 500	17 500	19 000	20 000	20 001	20 001	20 002	20 002
Reality	28 344	15 648	16 889	10 293	-1 297	21 661	15 977	20 064	

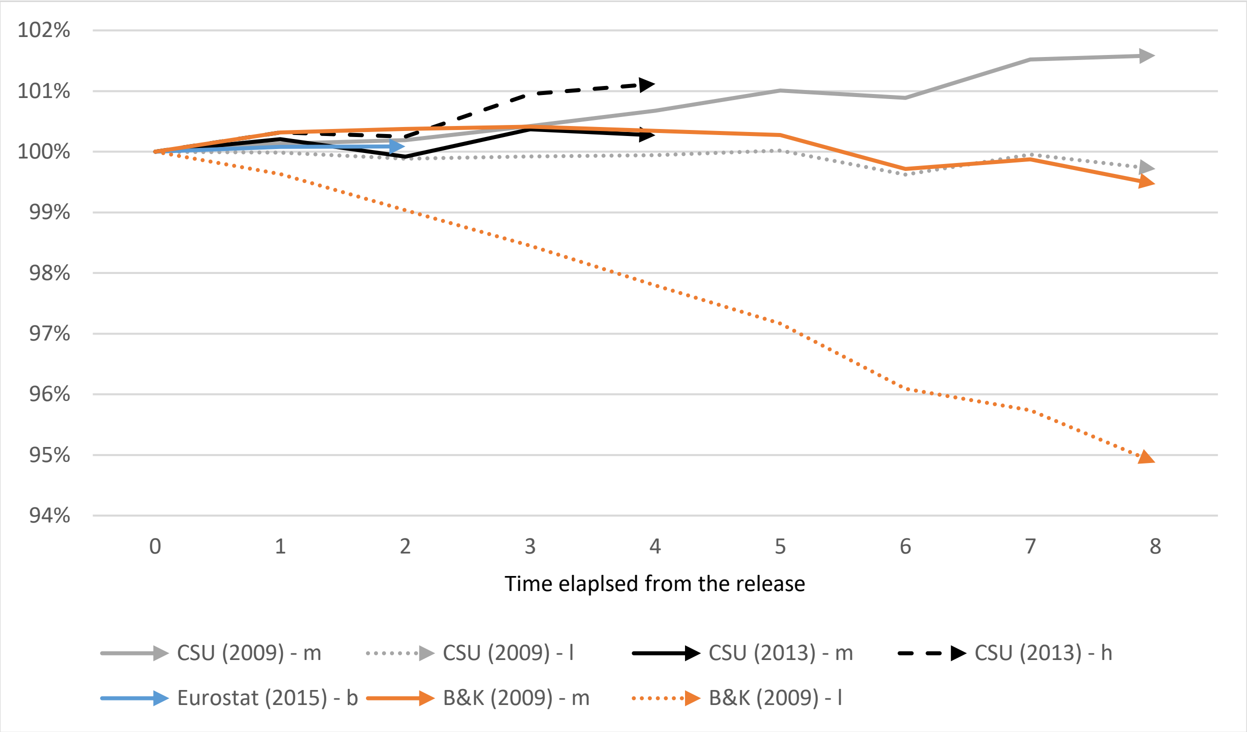
Source: CSU, Eurostat, B&K

9.4 Comparison of age groups

9.4.1 Old age groups

The accuracy of the projections is decreasing with increasing of the age of old-old (80+ years) part of the population. It is getting worse for 90+ years and even worse for 100+ years. The deviation from the reality reaches even 50%. Again, for the female part the accuracy is little bit higher rather than for the male part because of more observations in the female group. This tendency of lower accuracy of older age groups is observed for all the projections and mainly can be explained by the fact that there are getting less and less people in elder age groups, thus it is more and more difficult to be as precise as the real values. In other words, there is a high error margin that is inherent to older age groups.

Figure 26 Keyfitz's index. Population aged 80-89 years according to the individual projections and the time elapsed from the projections' release

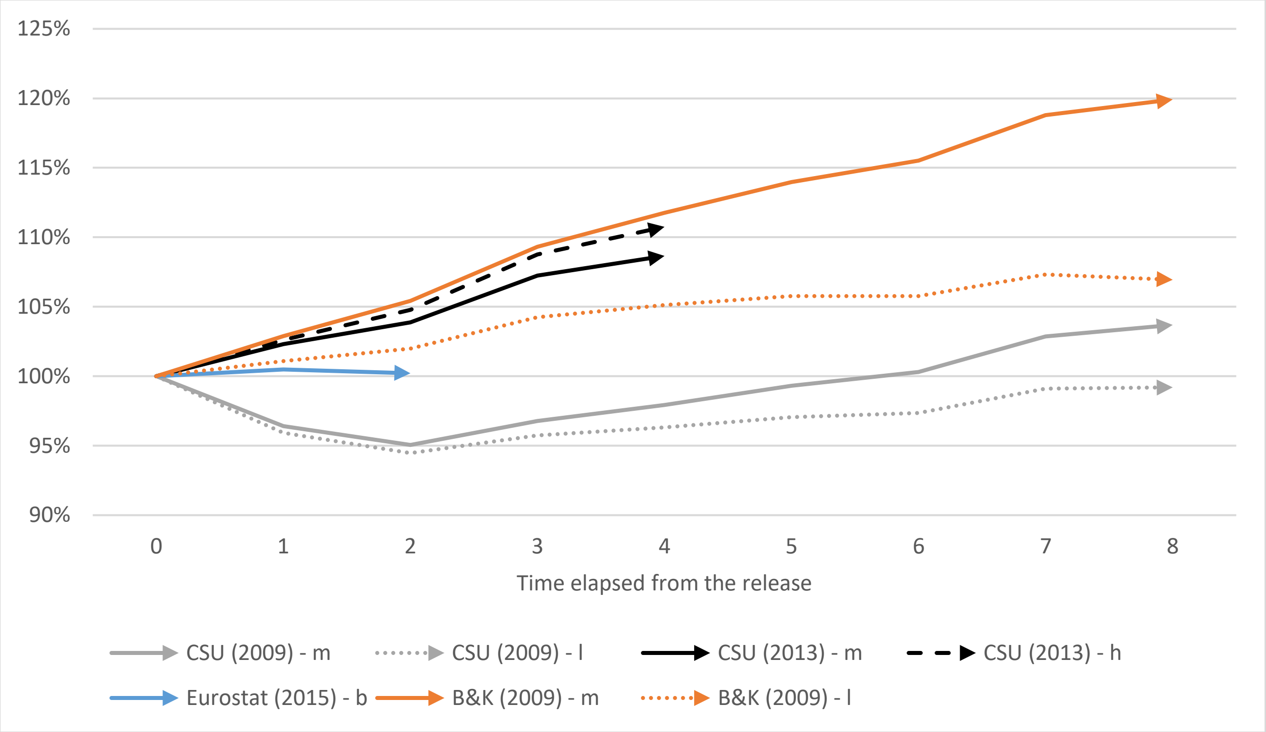


Source: CSU, Eurostat, B&K, own calculations

The graph 26 represents the results of the Keyfitz's index for the age group 80-89 years. In fact, the deviation does not seem that high for the most projections and belongs to the interval of [99.47%; 101.59%]. There is only one outlier in the graph, which is the low variant of B&K-2009 projection that is underestimated by more than 5% in the latest available year. According to MAPE, the average error for the whole period reaches 2.6% for the worst projection.

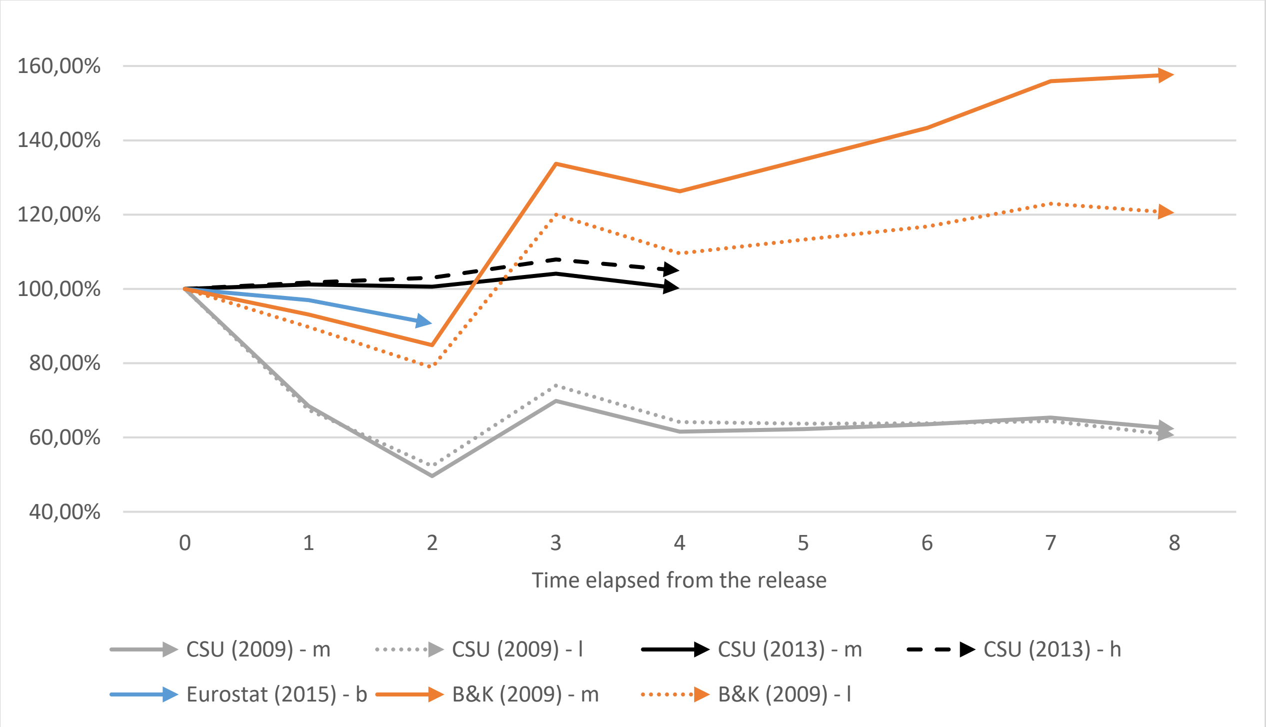
For the age group 90-99 years we can notice the evidence of the deviation to rise both up and down (see figure 27), the Keyfitz's index approximately belongs to the interval of [95%; 120%] saying that there is the rise of the deviation comparing to younger age groups. The deviation of the least accurate variant, B&K-2009 medium, reaches the deviation of 20% at the upper level. CSU-2009 both variants, B&K-2009 low variant and Eurostat-2005 performs with the relatively low deviation. According to MAPE, the average error for the whole period reaches 12.2% for the worst projection.

Figure 27 Keyfitz's index. Population aged 90-99 years according to the individual projections and the time elapsed from the projections' release



Source: CSU, Eurostat, B&K, own calculations

Figure 28 Keyfitz's index. Population aged 100+ years according to the individual projections and the time elapsed from the projections' release



Source: CSU, Eurostat, B&K, own calculations

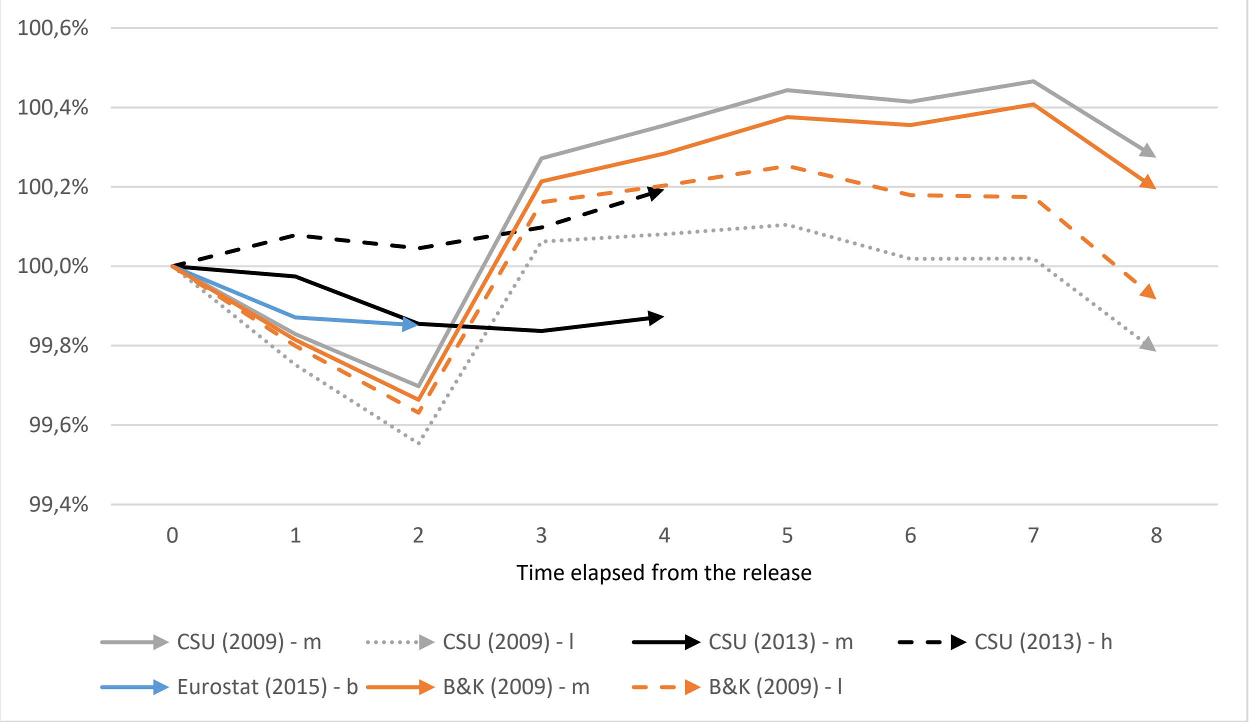
The same divergent trend is true for the age category 100+ years (see figure 28). This age category has the highest error margin, just few people can live up to this age and more, thus it is the most difficult age group to predict. On the figure 28 we can see that the deviation rises up to 57.69% at the upper level for the B&K-2009 – medium variant projection, and almost reaches 40% at the lower limit for both variants of CSU-2009. According to MAPE, the average error for the whole period reaches 37.1% for the worst projection.

Due to the high error margin and higher deviations of the projected values from the real values, the author of this work has elaborated special criteria ranges for age groups 80-89; 90-99 and 100+ years allowing wider interval boundaries to be considered as “good” accuracy. Such intervals differ for each older age group and can be found precisely in the sub-chapter chapter 5.4.

9.4.2 Most accurate age groups

The most accurate age categories belong to the ages from 10 to 19 and from 60 to 69. Almost all individual variants of all the projections for the individual years have got the assessment “good” according to the criteria listed in the table above. For example, on the figure 29 we can notice that the Keyfitz’s coefficient for the age group 10-19 years for all the individual projections belongs to the boundaries [99.55 %; 100.47 %], which means that the deviation does not exceed 0.47 % of the higher boundary and 0.45% of the lower boundary for any of the individual projections. According to MAPE, the average error for the whole period does not exceed 0.4% for the worst projection.

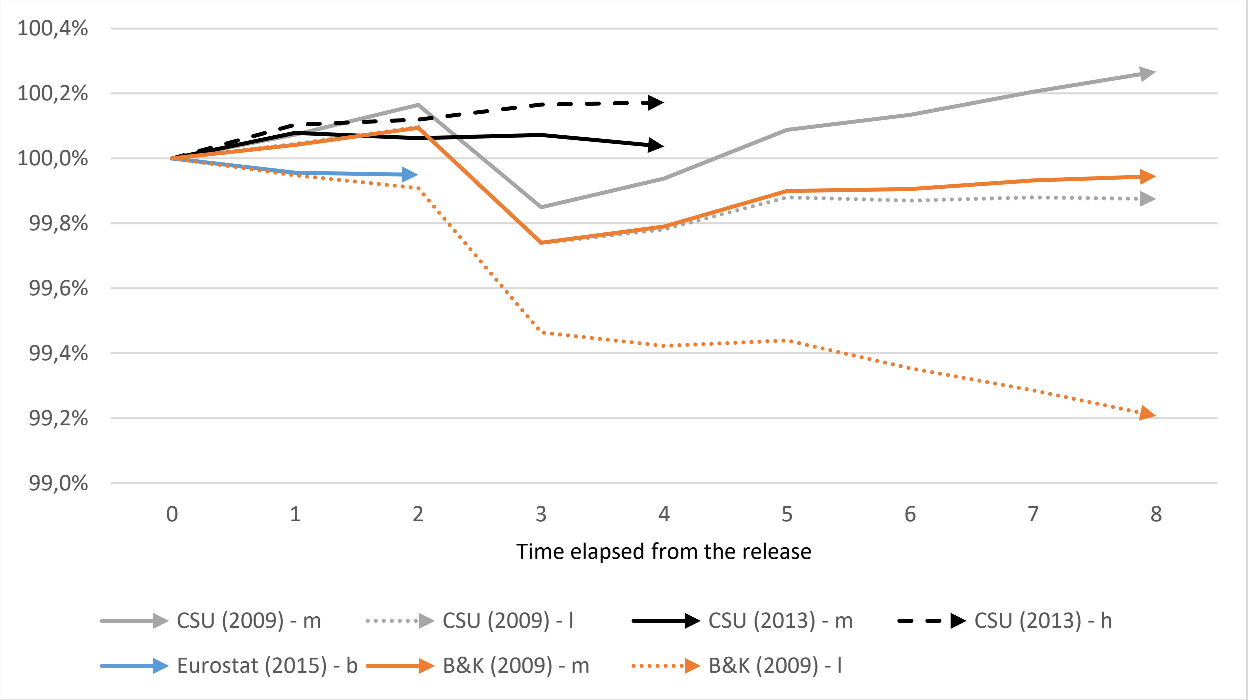
Figure 29 Keyfitz’s index. Population aged 10-19 years according to the individual projections and the time elapsed from the projections' release



Source: CSU, Eurostat, B&K, own calculations

Analogically, on the figure 30 we can see that the Keyfitz's coefficient for the age group 60-69 years for all the individual projections belongs to the boundaries [99.21 %; 100.27 %], which means that the deviation does not exceed 0.79 % of the lower boundary and 0.27% of the higher boundary for any of the individual projections. According to MAPE, the average error for the whole period does not exceed 0.5% for the worst projection.

Figure 30 Keyfitz's index. Population aged 60-69 years according to the individual projections and the time elapsed from the projections' release



Source: CSU, Eurostat, B&K, own calculations

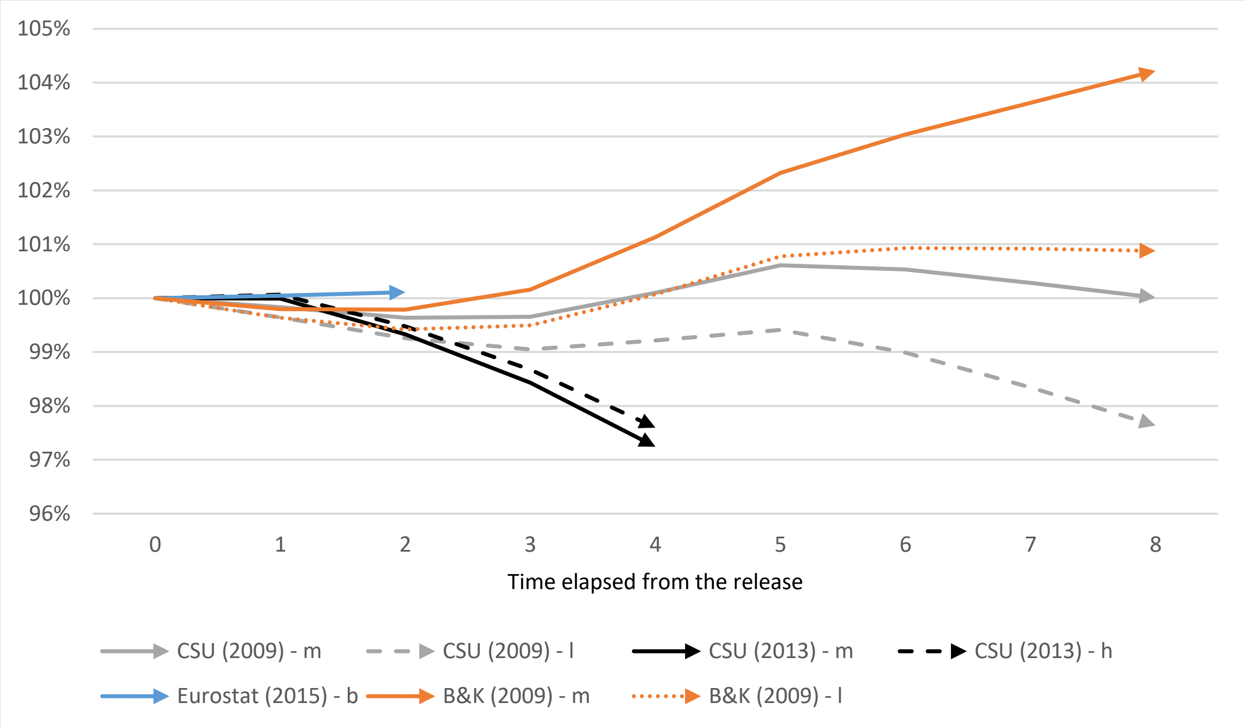
9.4.3 Least accurate age groups

Despite the old age categories there are also higher deviations for the age categories from 0 to 9 and from 20 to 39. The deviation fluctuates almost up to 5 %, and this can be clearly seen on the example of B&K-2009 – medium variant projection (figure 10), where the least accurate age groups are 0-9, 20-29, and 30-39 (despite the age group 90-99) and they are lying aside from other age groups in the graph.

If we look at the figure 36, we can notice the strong divergence of the individual projections for the age group from 0 to 9 years. Beside the older age groups (from 80 to 100+ years) this age group is the least accurate one. The Keyfitz's index for the ages 0-9 years for all the individual projections belongs to the boundaries [97.24 %; 104.23 %], which means that the deviation of the lower boundary is 2.76 % and 4.23% of the higher boundary, which is considerably wider comparing to other age groups (despite older age groups). According to MAPE, the average error for the whole period reaches 1.9% for the worst projection.

This can be explained by the complexity of prediction the number of births in the country, since total fertility rate is fluctuating, and age-specific fertilities and not easy to be set as a projection's assumption. Even small change in fertility rate, especially for bigger populations, will affect the number of births a lot.

Figure 31 Keyfitz's index. Population aged 0-9 years according to the individual projections and the time elapsed from the projections' release



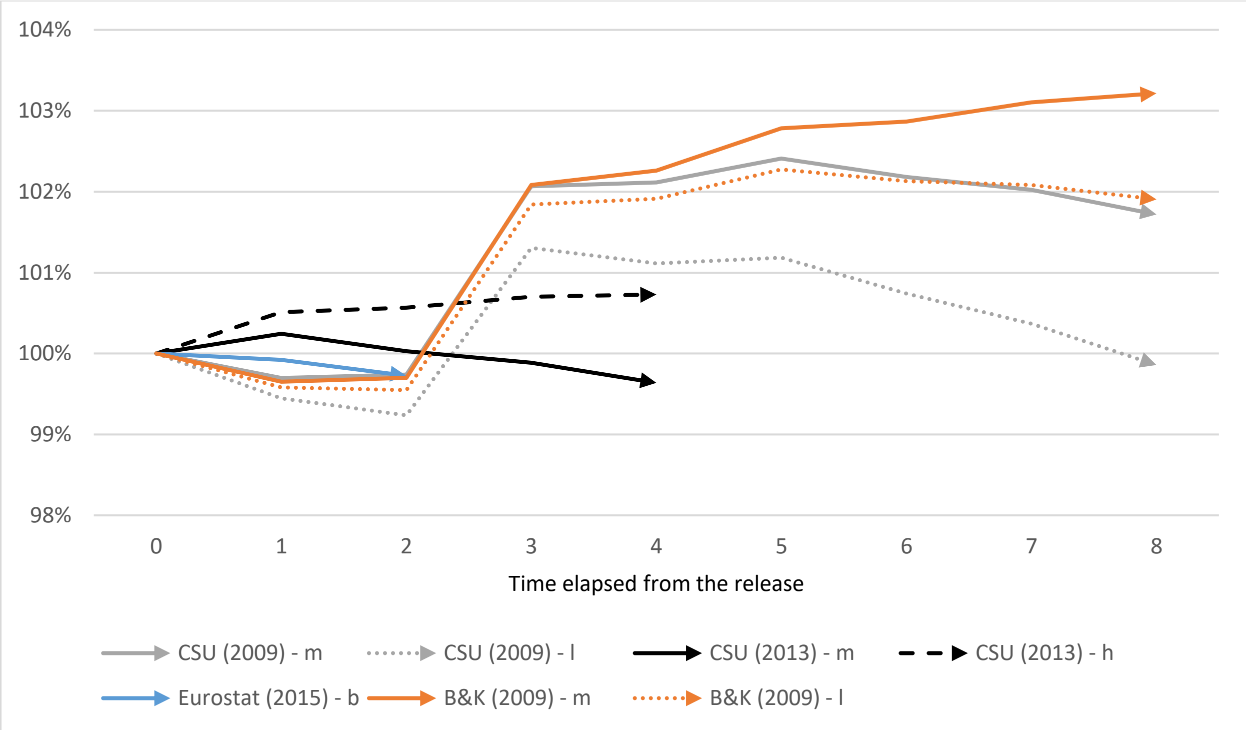
Source: CSU, Eurostat, B&K, own calculations

The below figure 32 represents the results of the Keyfitz's index of the second least accurate age group from 20 to 29 years. From the graph it is clearly seen that the most "successful" projections are Eurostat-2015 and the medium and the high variants of CSU-2013. The rest demonstrate poorer results with the medium variant of B&K, which shows the worst result among others for these ages, reaching the deviation of 3.22% for the latest year 2017. The Keyfitz's index belongs to the interval [99.24%; 103.22%] for the whole period, the projections usually are overestimated comparing to the reality. According to MAPE, the average error for the whole period reaches 2.1% for the worst projection.

The figure 33 demonstrates the results of the Keyfitz's index of the age group 30-39 years, which is in fact the third age group with lowest accuracy. Very similar to the ages 20-29, Eurostat-2015 and the both variants of CSU-2013 perform best of all, and the rest have higher deviation, such that the medium variants of B&K-2009 and CSU-2009 are deviated from the observed values by around 3.5%, and the whole Keyfitz belongs to the interval of [99.63%; 103.54%], which is almost the same as on the previous graph. According to MAPE, the average error for the whole period reaches 2.0% for the worst projection.

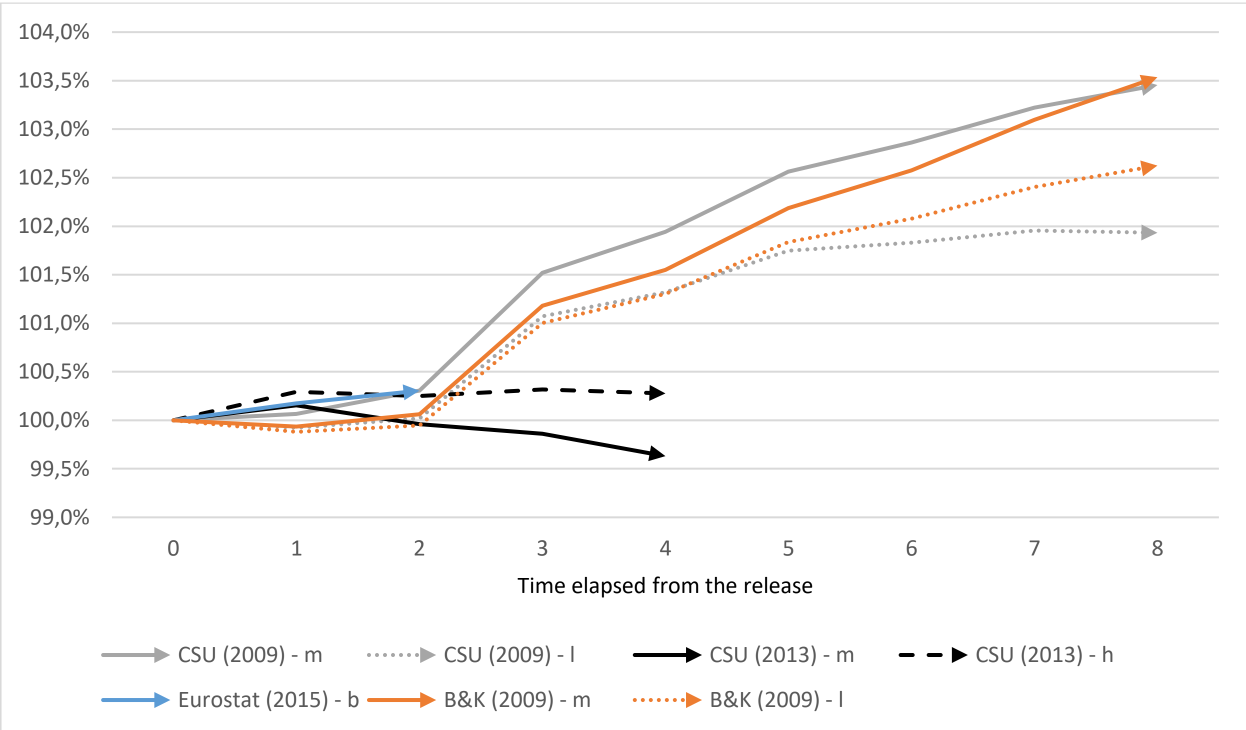
Probably such a high deviation of the ages from 20 to 39 can be explained by the fact that it is the most probable ages to emigrate and immigrate to the Czech Republic. According to Eurostat database, the most frequent age to leave the country belongs exactly to the interval from 20 to 39 years. Although the number of emigrants had been fallen through the years 2009-2016, the share of the emigrants aged from 20 to 39 years was prevailing in each (73% in 2009, 51% in 2016). Since all the projection performing higher deviation considerably overestimate the reality, we can make the conclusion that those producers did not expect that the outflow from the Czech Republic composed of this age group would be so high.

Figure 32 Keyfitz's index. Population aged 20-29 years according to the individual projections and the time elapsed from the projections' release



Source: CSU, Eurostat, B&K, own calculations

Figure 33 Keyfitz's index. Population aged 30-39 years according to the individual projections and the time elapsed from the projections' release



Source: CSU, Eurostat, B&K, own calculations

9.5 Overall results of Keyfitz and MAPE

All the main results are contained in the tables below that present the coefficients of the indicators and parameters of the individual projections according to the Keyfitz's coefficient by the weighted age categories and the MAPE method.

The results of both methods correspond to each other throughout the all stages of calculations. This fact makes the results sustainable between each other. For example, the table 15 shows that CSU-2009 – medium variant has less accurate results after the year 2013 for the male population weighted by age groups 0-14; 15-64; 65+, whereas MAPE (table 16), which shows the average deviation for the whole period, confirms it twice: firstly by proving the bad accuracy for the whole male population (1.27 %), and secondly getting higher deviation for the males at the age from 15 to 64 (1.5%).

If we look at the same Keyfitz's coefficient but for female part (again weighted by age groups 0-14; 15-64; 65+, CSU-2009 – medium variant), we can see that the results are quite stable and the coefficients of all the years do not belong to the bad region of >=1%. MAPE corresponds to Keyfitz, and the deviation does not exceed 1 % neither for the total female population, nor for any of the age groups (female 0-14; 15-64; 65+).

9.5.1 Keyfitz's index weighted by age groups

The table 15 with the Keyfitz's index weighted by the age groups is presented in the way that the forecasts to be comparable between each other. The errors of the projections are introduced in the table according to the time elapsed from their publications, such that it makes it possible to find what is the most accurate and the least accurate projection in each year.

The first thing that is evident is that the old projections (CSU-2009, B&K-2009) are starting to have a bigger deviation on average after the fourth year from the release. This can be explained by the fact that those projections are the oldest among others and there are more years spent from the projections' release. However, there are some exceptions like the low variant of CSU-2009 whose deviation is increasing as well but not that critical comparing to other variants of 2009 release. Also, both variants of CSU-2009 demonstrate good performance for the female part of the population for the both weighted indexes, such that the deviation does not exceed 1% even during the eighth year after the release.

As for the third and fourth year after the publication where the oldest projections can be also compared with the projection of CSU-2013, we can see that the newer projection demonstrates better accuracy especially for the Keyfitz's index weighted by the 10 years step interval. However, it is hard to say which variant of CSU-2013 is better since the lowest error is exchanged alternatively by the medium and high variant for the different years and populations.

Finally, analyzing the errors during the first and second year after the projections' release, we can see that the projection of Eurostat is the most accurate among others for the males, females as well as the total population measured by the indexes weighted by the both ways. Also, the good performance belongs to the medium variant of CSU-2013 for females, medium variant of CSU-2009 males and total population, medium variant of B&K-2009 for the all populations (male, female, total).

Table 15 Keyfitz's index weighted by age categories. Male and female population according to the individual projections and the time elapsed from the projections' release

Q_a	Forecast	0	1	2	3	4	5	6	7	8
Males (0-14;15-64;65+)	CSU (2009) - m	0,00%	0,07%	0,32%	0,88%	1,17%	1,59%	1,68%	1,84%	1,89%
	CSU (2009) - l	0,00%	0,11%	0,12%	0,63%	0,67%	0,82%	0,81%	0,91%	0,98%
	B&K (2009) - m	0,00%	0,11%	0,07%	0,60%	0,82%	1,25%	1,38%	1,61%	1,74%
	B&K (2009) - l	0,00%	0,20%	0,22%	0,56%	0,65%	0,95%	0,97%	1,00%	0,95%
	CSU (2013) - m	0,00%	0,13%	0,17%	0,41%	0,72%				
	CSU (2013) - h	0,00%	0,26%	0,27%	0,44%	0,58%				
	Eurostat (2015) - b	0,00%	0,03%	0,08%						
Females (0-14;15-64;65+)	CSU (2009) - m	0,00%	0,11%	0,15%	0,52%	0,57%	0,80%	0,80%	0,89%	0,92%
	CSU (2009) - l	0,00%	0,21%	0,35%	0,44%	0,45%	0,53%	0,54%	0,55%	0,63%
	B&K (2009) - m	0,00%	0,08%	0,06%	0,55%	0,78%	1,17%	1,32%	1,57%	1,73%
	B&K (2009) - l	0,00%	0,17%	0,27%	0,64%	0,76%	1,06%	1,18%	1,27%	1,35%
	CSU (2013) - m	0,00%	0,09%	0,09%	0,24%	0,42%				
	CSU (2013) - h	0,00%	0,18%	0,20%	0,38%	0,52%				
	Eurostat (2015) - b	0,00%	0,02%	0,04%						
Total (0-14;15-64;65+)	CSU (2009) - m	0,00%	0,03%	0,14%	0,69%	0,85%	1,19%	1,23%	1,36%	1,38%
	CSU (2009) - l	0,00%	0,15%	0,20%	0,53%	0,56%	0,67%	0,67%	0,69%	0,76%
	B&K (2009) - m	0,00%	0,10%	0,05%	0,57%	0,80%	1,20%	1,35%	1,59%	1,73%
	B&K (2009) - l	0,00%	0,19%	0,25%	0,60%	0,71%	1,01%	1,08%	1,13%	1,15%
	CSU (2013) - m	0,00%	0,11%	0,12%	0,32%	0,57%				
	CSU (2013) - h	0,00%	0,22%	0,23%	0,41%	0,55%				
	Eurostat (2015) - b	0,00%	0,02%	0,05%						
Males (10 years interval)	CSU (2009) - m	0,00%	0,17%	0,38%	1,01%	1,22%	1,60%	1,68%	1,84%	1,89%
	CSU (2009) - l	0,00%	0,19%	0,37%	0,76%	0,80%	0,96%	0,95%	1,03%	1,08%
	B&K (2009) - m	0,00%	0,17%	0,25%	0,81%	1,02%	1,40%	1,52%	1,71%	1,83%
	B&K (2009) - l	0,00%	0,21%	0,30%	0,84%	0,93%	1,20%	1,27%	1,32%	1,37%
	CSU (2013) - m	0,00%	0,14%	0,19%	0,40%	0,69%				
	CSU (2013) - h	0,00%	0,26%	0,28%	0,47%	0,66%				
	Eurostat (2015) - b	0,00%	0,08%	0,16%						
Females (10 years interval)	CSU (2009) - m	0,00%	0,12%	0,22%	0,58%	0,63%	0,85%	0,84%	0,91%	0,97%
	CSU (2009) - l	0,00%	0,21%	0,35%	0,48%	0,49%	0,57%	0,60%	0,69%	0,84%
	B&K (2009) - m	0,00%	0,12%	0,19%	0,66%	0,90%	1,28%	1,50%	1,74%	1,94%
	B&K (2009) - l	0,00%	0,18%	0,29%	0,74%	0,85%	1,13%	1,26%	1,37%	1,44%
	CSU (2013) - m	0,00%	0,13%	0,19%	0,32%	0,51%				
	CSU (2013) - h	0,00%	0,19%	0,26%	0,42%	0,56%				
	Eurostat (2015) - b	0,00%	0,08%	0,14%						
Total (10 years interval)	CSU (2009) - m	0,00%	0,14%	0,29%	0,77%	0,88%	1,19%	1,22%	1,33%	1,34%
	CSU (2009) - l	0,00%	0,18%	0,31%	0,59%	0,62%	0,73%	0,74%	0,77%	0,85%
	B&K (2009) - m	0,00%	0,14%	0,22%	0,71%	0,95%	1,33%	1,50%	1,70%	1,85%
	B&K (2009) - l	0,00%	0,19%	0,29%	0,77%	0,85%	1,14%	1,22%	1,30%	1,34%
	CSU (2013) - m	0,00%	0,12%	0,14%	0,32%	0,55%				
	CSU (2013) - h	0,00%	0,22%	0,25%	0,43%	0,58%				
	Eurostat (2015) - b	0,00%	0,07%	0,15%						

Source: own calculations, CSU, Eurostat, B&K

The summary of the results presented in the table 15 is contained in the table 16. In particular, this table reflects the most and least accurate projections for each year after the projections' publication for the total population. On average, we can see that during the first two years the projection of Eurostat leads, after the position of the most accurate projection is held by the medium variant of CSU-2013, and after the fourth year the low variant of CSU-2009 demonstrates the best result. As for the worst projections, for the first four years we can see the presence of the low variant of B&K-2009, medium and low variant of CSU-2009, high variant of CSU-2013. After the fourth year from the publication it is clearly seen that the medium variant of B&K performs worst of all.

Table 16 Keyfitz's index weighted by age categories. Best and worst projections according to the time elapsed from the projections' release

Qa	Time	1	2	3	4	5	6	7	8
Total (0-14;15-64;65+)	*Best*	Eurostat 0.02%	Eurostat 0.05%	CSU-13 m 0.32%	CSU-13 h 0.55%	CSU-09 l 0.67%	CSU-09 l 0.67%	CSU-09 l 0.69%	B&K-09 l 0.76%
	Worst	B&K-09 l 0.19%	B&K-09 l 0.25%	CSU-09 m 0.69%	CSU-09 m 0.85%	B&K-09 m 1.20%	B&K-09 m 1.35%	B&K-09 m 1.59%	B&K-09 m 1.73%
Total (10 years interval)	*Best*	Eurostat 0.07%	CSU-13 m 0.14%	CSU-13 m 0.32%	CSU-13 m 0.55%	CSU-09 l 0.73%	CSU-09 l 0.74%	CSU-09 l 0.77%	CSU-09 l 0.85%
	Worst	CSU-13 h 0.22%	CSU-09 l 0.31%	B&K-09 l 0.77%	B&K-09 m 0.95%	B&K-09 m 1.33%	B&K-09 m 1.50%	B&K-09 m 1.70%	B&K-09 m 1.85%

Source: own calculations, CSU, Eurostat, B&K

9.5.2 MAPE

The table 17 demonstrates the results of MAPE, in other words it shows the average deviation of the projected values from the real vales for the whole period. With the help of this table we can promptly see what the problematic areas (red cells) are and the best results (green cells) for each individual projection. For example, all the projections have the bad results for the male life expectancy at 65. Almost the same the situation is with the female life expectancy at 65, despite the low variant of CSU-2009, which performs with the average error of 0.81% for the whole period of 8 years. Also, it is clearly visible how huge is the deviation of projected net migration reaching even 399.5% in case of the high variant of CSU-2013.

The author would like to stress out that the results of MAPE presented in the table 17 are intended not for the comparison of the projections but for the representation of the average error performed by the individual projections. The error cannot be compared because of the different publication year of the projections and because of the different time period involved in the calculation which can be observed at the top of the table separate for each projection. From this table the individual performance of the projections can be understood.

However, something can still be comparable. First of all, the performance of different variants of the same projection can be traced. Secondly, the older projections can be compared between each other since MAPE was calculated for the same period of time.

As for CSU-2009, we can see that the medium variant has more deviated projection in comparison to the low variant. This is indicated by the higher number of the red cells reflecting the high deviation. The low variant has one green cell which is an indicator of the very accurate performance, in particular, the average error of the projected female population aged 60-69 years does not exceed 1% for the whole period of 8 years and equals to 0.041%.

As for B&K-2009, there is no clear pattern of one variant to be better or worse than another. Both variants perform with the higher errors for the whole period, and just the medium variant demonstrates very good result for the female population aged 60-69 years with the average error of 0.056% for the whole period of 8 years.

As for CSU-2013, it can be seen that the medium variant performs better than the high because of the more components projected very accurate (green cells) by the medium variant and more components projected very inaccurate (red cells) by the high variant. MAPE is calculated for the period of 5 years.

Table 17 MAPE. Individual components according to the individual projections and the number of years involved

Number of years involved	8	8	8	8	5	5	2
MAPE	CSU (2009) - m	CSU (2009) - l	B&K (2009) - m	B&K (2009) - l	CSU (2013) - m	CSU (2013) - h	Eurostat (2015) - b
Total population	0,896%	0,301%	0,965%	0,357%	0,198%	0,139%	0,001%
Males	1,272%	0,543%	1,013%	0,327%	0,216%	0,179%	0,028%
Females	0,621%	0,128%	0,984%	0,399%	0,180%	0,100%	0,026%
0-14 males	0,440%	0,507%	1,475%	0,555%	0,825%	0,672%	0,058%
0-14 females	0,269%	0,841%	1,383%	0,481%	0,911%	0,756%	0,060%
15-64 males	1,500%	0,763%	0,982%	0,597%	0,212%	0,241%	0,057%
15-64 females	0,793%	0,399%	1,066%	0,819%	0,084%	0,229%	0,006%
65+ males	0,470%	0,133%	0,217%	1,263%	0,519%	0,721%	0,044%
65+ females	0,193%	0,388%	0,063%	1,162%	0,096%	0,287%	0,061%
80+ males	1,920%	0,982%	2,458%	1,264%	1,459%	2,001%	0,097%
80+ females	0,401%	1,037%	1,225%	1,956%	0,572%	1,147%	0,210%
0-9 males	0,448%	0,819%	1,992%	0,725%	1,187%	1,018%	0,084%
0-9 females	0,299%	1,307%	1,732%	0,549%	1,341%	1,145%	0,075%
10-19 males	0,360%	0,159%	0,194%	0,153%	0,056%	0,169%	0,171%
10-19 females	0,313%	0,160%	0,412%	0,293%	0,185%	0,060%	0,105%
20-29 males	2,222%	1,191%	2,358%	1,811%	0,170%	0,768%	0,218%
20-29 females	1,020%	0,510%	1,871%	1,433%	0,218%	0,482%	0,131%
30-39 males	2,525%	1,572%	1,790%	1,374%	0,253%	0,301%	0,172%
30-39 females	1,432%	0,929%	1,773%	1,485%	0,112%	0,266%	0,311%
40-49 males	1,611%	0,881%	0,621%	0,265%	0,296%	0,118%	0,042%
40-49 females	0,844%	0,501%	0,931%	0,713%	0,067%	0,232%	0,030%
50-59 males	0,713%	0,267%	0,187%	0,281%	0,223%	0,075%	0,096%
50-59 females	0,441%	0,209%	0,519%	0,336%	0,066%	0,167%	0,063%
60-69 males	0,183%	0,258%	0,263%	0,768%	0,078%	0,180%	0,066%
60-69 females	0,116%	0,041%	0,056%	0,256%	0,049%	0,106%	0,030%
70-79 males	0,308%	0,214%	0,494%	1,737%	0,389%	0,547%	0,107%
70-79 females	0,111%	0,204%	0,554%	1,284%	0,234%	0,108%	0,034%
80-89 males	1,519%	0,680%	0,697%	2,467%	0,898%	1,365%	0,077%
80-89 females	0,452%	0,526%	0,388%	2,745%	0,210%	0,305%	0,169%
90-99 males	5,793%	4,028%	19,606%	11,625%	7,034%	8,262%	0,069%
90-99 females	3,387%	5,176%	9,807%	2,568%	5,035%	6,224%	0,495%
100+ males	42,954%	42,446%	41,094%	24,833%	11,538%	9,904%	17,046%
100+ females	35,391%	34,362%	32,314%	14,551%	5,455%	8,695%	3,019%
Total fertility rate	3,038%	2,828%	6,141%	3,032%	4,649%	4,582%	2,135%
Net migration	298,311%	178,344%	309,969%	226,548%	230,045%	392,466%	3,665%
Life expectancy at birth males	0,294%	0,274%	0,169%	1,122%	0,204%	0,435%	0,370%
Life expectancy at birth females	0,359%	0,214%	0,220%	0,809%	0,226%	0,449%	0,405%
Life expectancy at 65 males	2,745%	1,473%	1,979%	1,909%	1,061%	2,639%	1,270%
Life expectancy at 65 females	1,648%	0,811%	1,112%	2,187%	1,026%	1,497%	1,428%
0-9 total	0,302%	1,057%	1,865%	0,628%	1,251%	1,080%	0,080%
10-19 total	0,337%	0,149%	0,294%	0,203%	0,115%	0,104%	0,139%
20-29 total	1,635%	0,771%	2,121%	1,627%	0,187%	0,628%	0,175%
30-39 total	1,992%	1,244%	1,782%	1,428%	0,176%	0,284%	0,239%
40-49 total	1,238%	0,681%	0,772%	0,484%	0,155%	0,149%	0,036%
50-59 total	1,483%	0,967%	1,265%	1,227%	0,923%	1,157%	0,157%
60-69 total	0,143%	0,139%	0,115%	0,496%	0,063%	0,141%	0,047%
70-79 total	0,187%	0,177%	0,528%	1,473%	0,066%	0,168%	0,062%
80-89 total	0,803%	0,125%	0,333%	2,652%	0,232%	0,659%	0,084%
90-99 total	2,674%	3,105%	12,201%	4,779%	5,528%	6,727%	0,355%
100+ total	37,130%	36,215%	34,212%	16,790%	1,514%	4,384%	6,177%
0-14 total	0,341%	0,670%	1,430%	0,519%	0,865%	0,713%	0,059%
15-64 total	1,133%	0,580%	1,020%	0,706%	0,149%	0,235%	0,026%
65+ total	0,257%	0,208%	0,117%	1,204%	0,257%	0,467%	0,038%

Source: own calculations, CSU, Eurostat, B&K

Table 18 MAPE. Best and worst projections for the whole period (CSU-2009 and B&K-2009)

Component	MAPE	
	Best for the period of 8 years	*Worst for the period of 8 years*
Total population	CSU – low, 0.301%	CSU – medium, 0.896%
Males	B&K – low, 0.321%	CSU – medium, 1.272%
Females	CSU – low, 0.128%	B&K – medium, 0.984%
0-9 males	CSU – medium, 0.448%	B&K – medium, 1.992%
0-9 females	CSU – medium, 0.299%	B&K – medium, 1.732%
0-9 total	CSU – medium, 0.302%	B&K – medium, 1.865%
10-19 males	B&K – low, 0.153%	CSU – medium, 0.360%
10-19 females	CSU – low, 0.160%	B&K – medium, 0.412%
10-19 total	CSU – low, 0.149%	B&K – medium, 0.294%
20-29 males	CSU – low, 1.191%	B&K – medium, 2.358%
20-29 females	CSU – low, 0.510%	B&K – medium, 1.871%
20-29 total	CSU – low, 0.771%	B&K – medium, 2.121%
30-39 males	B&K – low, 1.374%	CSU – medium, 2.525%
30-39 females	CSU – low, 0.929%	B&K – medium, 1.773%
30-39 total	CSU – low, 1.244%	CSU – medium, 1.992%
40-49 males	B&K – low, 0.265%	CSU – medium, 1.611%
40-49 females	CSU – low, 0.501%	B&K – medium, 0.931%
40-49 total	B&K – low, 0.484%	CSU – medium, 1.238%
50-59 males	B&K – medium, 0.187%	CSU – medium, 0.713%
50-59 females	CSU – low, 0.209%	B&K – medium, 0.519%
50-59 total	CSU – low, 0.967%	CSU – medium, 1.483%
60-69 males	CSU – medium, 0.183%	B&K – low, 0.768%
60-69 females	CSU – low, 0.041%	B&K – low, 0.256%
60-69 total	B&K – medium, 0.115%	B&K – low, 0.496%
70-79 males	CSU – low, 0.214%	B&K – low, 1.737%
70-79 females	CSU – medium, 0.111%	B&K – low, 1.284%
70-79 total	CSU – low, 0.177%	B&K – low, 1.473%
80-89 males	CSU – low, 0.680%	B&K – low, 2.467%
80-89 females	B&K – medium, 0.388%	B&K – low, 2.745%
80-89 total	CSU – low, 0.125%	B&K – low, 2.652%
90-99 males	CSU – low, 4.028%	B&K – medium, 19.606%
90-99 females	B&K – low, 2.568%	B&K – medium, 9.807%
90-99 total	CSU – medium, 2.674%	B&K – medium, 12.201%
100+ males	B&K – low, 24.833%	CSU – medium, 42.954%
100+ females	B&K – low, 14.551%	CSU – medium, 35.391%
100+ total	B&K – low, 16.790%	CSU – medium, 37.130%
Total fertility rate	CSU – low, 2.828%	B&K – medium, 6.141%
Net migration	CSU – low, 178.344%	B&K – medium, 309.969%
Life expectancy at birth males	B&K – medium, 0.169%	B&K – low, 1.122%
Life expectancy at birth females	CSU – low, 0.214%	B&K – low, 0.809%
Life expectancy at 65 males	CSU – low, 1.473%	B&K – medium, 1.979%
Life expectancy at 65 females	CSU – low, 0.811%	B&K – low, 2.187%

Source: own calculations, CSU, B&K

Being that the projection CSU-2009 and B&K-2009 are comparable between each other, it makes sense to check which is the most accurate and the least accurate one for the individual components of the population. The summary of the comparison is contained in the table 18. It can be said that on average the projection CSU performs better than the projection of B&K. However, there are some exceptions: for example, the low variant of B&K shows very accurate results for the people aged 100+. On the contrary, the medium variant of CSU performs worst of all for this age group.

9.5.3 Span of the results

The analysis has proven that the projections cannot be uniformly judged. If one variant performs better than others or at least it belongs to the category of "good" accuracy for some age groups or parameters that does not necessarily mean that it will hold the "good" assessment for other age groups and parameters. Some of the variants demonstrate very good performance to the particular categories but at the same time could be the worst to other categories. Also, the different calendar years of the projections' releases do not let them be comparable.

The results obtained by MAPE can be compared just between the projections released during the same year. To compare the projections by the results of the Keyfitz's index, the proper time elapsed from the projections' release must be chosen. All the projections can be compared just during the first 2 years in accordance with their dates of the publication. To be able to compare the projection of Eurostat and CSU-2013 with the older projections the author of this work decided to compare the results of Keyfitz for the second year after the release of each projection. In particular, it is done to present the deviation interval for this year, and to find the most accurate and the least accurate projection for the selected category. The integration is presented in the table 19.

From table 19 we can see that the deviation interval is quite narrow, and it can be said that the accuracy of all the projections is high during the second year after the publication, except for some problematic parameters. The projection of Eurostat is the most frequent projection to be the best among others. Moreover, its error never exceeds the bound to belong to the "worst" category throughout all the components of the population.

Also, the comparison proves that the low variant of CSU-2009 and the medium variant of CSU-2013 also perform with the higher accuracy, however in several cases those variants can be found among "worst" projections. Among the column "worst" most often we can see the medium variant of CSU-2009. We can see that during the first 2 years the projection of CSU-2009 does not perform very well, however for the whole period (table 18) it demonstrates better results. On the contrary, the projection of B&K-2009 performs with the average result during the first 2 years, however for the whole period the error grows.

Table 19 Keyfitz's index. Span of the results for the second year

Component	Keyfitz		
	Interval	*Best during the second year*	*Worst during the second year*
Total population	[99.92; 100.49]	Eurostat	CSU (2009) - m
Males	[99.93; 100.68]	Eurostat	CSU (2009) - m
Females	[99.91; 100.39]	Eurostat	B&K - m
0-9 total	[99.33; 100.11]	Eurostat	CSU (2009) - l
10-19 total	[99.55; 100.05]	CSU (2013) - h	CSU (2009) - l
20-29 total	[99.24; 100.57]	CSU (2009) - l	CSU (2013) - m
30-39 total	[99.95; 100.31]	CSU (2009) - l	CSU (2009) - m
40-49 total	[99.95; 100.43]	Eurostat	CSU (2009) - m
50-59 total	[99.90; 100.26]	B&K - l	CSU (2009) - m
60-69 total	[99.91; 100.16]	Eurostat	CSU (2009) - m
70-79 total	[99.16; 100.12]	CSU (2009) - m	B&K - l
80-89 total	[99.04; 100.38]	CSU (2013) - m	B&K - l
90-99 total	[94.47; 105.43]	Eurostat	CSU (2009) - l
100+ total	[49.60; 103.01]	CSU (2013) - m	CSU (2013) - h
Total fertility rate	[95.02; 103.14]	Eurostat	CSU (2013) - m
Net migration	[40.34; 159.76]	CSU (2009) - l	CSU (2009) - m
Life expectancy at birth males	[99.21; 100.51]	CSU (2013) - m	B&K - l
Life expectancy at birth females	[99.24; 100.28]	B&K - m	B&K - l
Life expectancy at 65 males	[98.64; 102.90]	Eurostat	CSU (2009) - m
Life expectancy at 65 females	[97.87; 101.02]	CSU (2009) - l	B&K - l

Source: own calculations, CSU, Eurostat, B&K

10 CONCLUSION AND DISCUSSION

The aim of the work was the post-evaluation of the accuracy of the individual population projections in the Czech Republic and comparison of the projections between each other. The aim was achieved with the help of 2 main methods, namely the Keyfitz's "Quality of prediction index" and the Mean Absolute Percentage Error. In total, 5 individual projections and 13 variants were used in the work from the 4 main sources: Czech statistical office, Eurostat, United Nations and the projection of Boris Burcin and Tomáš Kučera. The study has proven that the projections cannot be uniformly judged and compared between each other because of the different publication time. The results of Keyfitz can be relevantly compared just for the first 2 years after the projection's release because the newest projection of Eurostat was published in 2015. The outcomes of MAPE can be only compared in case the projections belong to the same calendar year or between the individual variants of a single projection. The two methods demonstrated the results in accordance to each other, so the results proved to be consistent.

According to the results of Keyfitz during the second year from the publication, Eurostat's projection with the baseline variant has the most accurate performance among others. The least accurate projection belongs to CSU-2009. However, with more time spent the situation changes, and after 2 years elapsed from the release the least accurate projection belongs to B&K. For the period during the third and fourth years CSU-2013 shows better results than CSU-2009 and B&K. For the period during the fourth to eighth year after the release where just CSU-2009 and B&K can be compared, CSU-2009 performs with more accurate results. The comparison of CSU-2009 and B&K with MAPE also confirmed what was said in the previous sentence. Also, MAPE has proven that the low variants of CSU-2009 and B&K demonstrate much better accuracy than their medium variants. As for CSU-2013, the high and the medium variants have pretty much the same performance but the medium one is still little bit more accurate.

The analysis has shown that with the time spent we can notice the divergence of the projections and their variants from the reality and within each other. In addition, this confirmed by the fact that the more years elapsed from the release of the projections, the larger the deviation from the real data is. According to the results of the study, the accuracy of the prediction seems to be very high during the first 2 years after the publication not exceeding the deviation of 1% for any of the projections except for some problematic components. The error starts to rise after 4 years elapsed from the projections' release exceeding the deviation of 1% and more. Another important notice is that the projections usually better constructed for the female part of the population regardless which age group it belongs. The newer projections (CSU 2013, Eurostat 2015) seem to have better results than the older ones (CSU 2009, B&K 2009) mainly thanks to their recent release and partly thanks to the developments and improvements based on the previous projections. Overestimation is typically observed for the older projections.

Regarding the age groups, the analysis has demonstrated that the accuracy of older age categories is decreasing with increasing the age. Not surprisingly that the age category 100+ is the least accurate, with the error even exceeding 50% during the 8th year after the publication, since it is hard to predict how many people will live up to 100 years and more, thus we are dealing with the high error margin. Also, the age category 0-9 years has the higher deviation from the reality, about 4% during the 8th year after the publication, because of the complexity of predicting the number of births. The theory confirms that the higher errors are inherent to the youngest and the oldest age groups. According to McDonald and Kippen (2008), "*in most advanced countries over the past 50 years, statistical agencies have performed poorly in projectioning the future number of births*". Also, it was confirmed by D. Lee (2012) that "*in post-transition populations there are no secular trends in fertility*". In addition, the analysis has proven that there are other groups in the projections of the Czech Republic that perform with the lower accuracy. They are the ages 20-39 with the deviation reaching 3.5%, and the errors are connection to the most popular age among the emigrants and immigrants which place more uncertainty to those ages. According to the results of Keyfitz and MAPE, the most accurate age groups are 10-19 and 60-69 with the deviation lower that 0.5%. The latter group (60-69) is especially important for the usage of the reform of the pension system in the Czech Republic. The most problematic parameters are net migration due to vague assumptions and individual factors to migrate that are hard to predict, and life expectancy at 65 due to the high error margin and changing future trends in mortalities which is absolutely corresponds to the theory.

Projecting the future populations play a big role at the governmental level since they are used for a wide range of planning and budgeting purposes. It is important to monitor how accurate the population projections are and to which extent they fit the real values, since many state and local decisions, which firstly influence the economy, are made based on the future population and future structural changes in the population size. The author of the thesis believe that it is necessary to check the accuracy of population projections not just to follow the tendencies of the projections and deviation from the observations but also to prevent further errors in future populations projections and actualize current sets of assumptions. It is important to start evaluating the accuracy at least after the 4th year from one's projection release since during this period the deviation may be rising and may have reached high values after the period of four years. After that period the author recommends monitoring the accuracy at least once per two years because the deviation may rise faster. Also, it is important to understand where the bias comes from: whether the reasons of poor-quality projections lies in incorrectly assumed vital rates, or the error comes from the bad collection of statistical estimates. Both reasons put some level of uncertainty on the results of population projections.

It is not possible to produce absolutely credible projection in the long-term perspective, however, there are some parts of the projections that can be improved by correcting relevant parameters based on the given analysis and trends in the population, like total fertility rate or life expectancy at particular ages. Also, there are some parts, like net migration or migration rate that can be hardly developed to obtain better results. It is important for users of projections and policymakers to understand how accurate the given projections are. Policymakers could employ population projections more effectively in the planning activities, if they would be more educated on the methodology of projections and their potential errors, weaknesses and uncertainty. Users could become more aware of actual reliability and restrictions of projected data.

The author finds both the Keyfitz's index and the MAPE method relevant for measuring the accuracy of population projections. The Keyfitz's coefficient represents the simple index where you judge the predicted value against the observed one demonstrating whether the projection is underestimated or overestimated. The higher variation between different age groups can be solved with the Keyfitz's index weighted by age groups. This index considers the accuracy of individual age groups and after weights them into a single index. The only disadvantage of Keyfitz weighted by age groups is that it is not possible to determine whether the projection is overestimated or underestimated. It is important to calculate the Keyfitz's index not for calendar years but for the time spent after the projections' publication while compare the projections between each other, hence, to be able to find the most accurate and the least accurate projection in each year.

As for MAPE, it can be highly useful when it is needed to calculate the average error for the given period. However, alike with the previous method (Keyfitz) it is impossible to judge whether the predictions are overvalued or not. MAPE has one big disadvantage of distorting the outcome if the calculation involves some outliers like when calculating any average value. It was checked on the example of net migration, when the deviation calculated by MAPE was extremely high for almost any of the projections. The Keyfitz's indices revealed the reason of such a poor MAPE results giving the information about the absolute discrepancy of predicted and observed values of net migration in 2013. Just one "bad" year can influence the outcome of MAPE a lot. While comparing the projections between each other with MAPE, it is important to understand what can be comparable since the outcomes of MAPE can be compared only if the projections were released in the same year.

As was mentioned in the introduction not so much work was done with the evaluation of population projections in the Czech Republic. So, with this project I would like to contribute to demographic statistics by providing the thorough analysis of the population projections and giving the base for further researches and developments. For example, knowing the results of the previous evaluation of the population projections together with the fresh results of the current projections and together with the methodological aspects of each population projection, the further research can be done in the field of the bias origin, depending either on the poor data collection or poor set of assumptions, and how the potential bias can be eliminated. I believe that the work detects the weakest and most problematic areas which will help to prevent the errors in further projections and improve them in future.

REFERENCES

American Planning Association (1950). *Population Projectioning.* Chicago, American Society of Planning Officials. Retrieved March 4, 2019 from https://www.planning.org/pas/reports/report17.htm

Australian Bureau of Statistics (2013). *Statistical Language - Estimate and Projection.* Retrieved February 13, 2019 from http://www.abs.gov.au/websitedbs/a3121120.nsf/home/statistical+language+-+estimate+and+projection

Billari F., Graziani R. & Melilli E. (2012). *Stochastic population projections based on conditional expert opinions.* Journal of the Royal Statistical Society. Series A (Statistics in Society). *175*(2):491-511. Retrieved September 17, 2018 from http://www.jstor.org/stable/23251247

Bull P. (1987). *Understanding population projections*. National Center for Biotechnology Information. *42*(4):1-43. Retrieved February 15, 2019 from https://www.ncbi.nlm.nih.gov/pubmed/12315299

Burcin B. & Kučera T. (2008). *Population projection 2008, Czech Republic*.

Burcin B. & Kučera T. (2010). *Prognóza populačního vývoje české republiky na období 2008–2070.* Ministerstvo práce a sociálních věcí. Retrieved February 26, 2019 from https://www.mpsv.cz/files/clanky/8842/Prognoza_2010.pdf

Capuano G. (2015). *On population projections, projections and weather.* Id. The population experts. Retrieved February 13, 2019 from https://blog.id.com.au/2015/population-projectioning/on-population-projections-projections-and-weather/

Caswell H. & Gassen N. S. (2015). *The sensitivity analysis of population projections.* Demographic Research. (33):801-840. Retrieved February 12, 2019 from http://dx.doi.org/10.4054/DemRes.2015.33.28

Chi G. (2009). *Can Knowledge Improve Population Projections at Subcounty Levels?* Demography. *46*(2):405-427. Retrieved from http://www.jstor.org/stable/20616470

Clements B. (2016). *The Absolute Best Way to Measure Projection Accuracy.* Retrieved August 15, 2018, from https://www.axsiumgroup.com/the-absolute-best-way-to-measure-projection-accuracy-2/

Czech Statistical Office (2018). *Czech Demographic Handbook 2017.* Publications of Czech Statistical Office. Retrieved March 18, 2019 from https://www.czso.cz/csu/czso/czech-demographic-handbook-2017

Czech Statistical Office. *Databases and registers*. Retrieved August 15, 2018, from https://www.czso.cz/csu/czso/databases-registers

Czech Statistical Office. *Demografie, Review for Population Research*. Retrieved November 10, 2018, from https://www.czso.cz/csu/czso/demografie_review_for_population_research

Czech Statistical Office (2018). *Population Change 2017.* Publications of Czech Statistical Office. Retrieved March 18, 2019 from https://www.czso.cz/csu/czso/ari/population-change-year-2017

Czech Statistical Office (2018). *Population Change 1st half of 2018.* Publications of Czech Statistical Office. Retrieved March 18, 2019 from https://www.czso.cz/csu/czso/ari/population-change-1st-half-of-2018

Czech Statistical Office (2009). *Population projection 2009, the Czech Republic*.

Czech Statistical Office (2013). *Population projection 2013, the Czech Republic.*

Czech Statistical Office (2018). *Projekce obyvatelstva České republiky 2018-2100.* Publications of Czech Statistical Office. Retrieved March 19, 2019 from https://www.czso.cz/documents/10180/61566242/13013918u.pdf/6e70728f-c460-4a82-b096-3e73776d0950?version=1.2

Dotson J. Dianne. (2018). *How to Calculate Population Projections.* Sciencing. Probability & Statistics. Retrieved February 15, 2019 from https://sciencing.com/calculate-population-projections-8473012.html

Environmental Systems Research Institute (2007). *Evaluating Population Projections - The Importance of Accurate Projectioning.* ESRI library. Retrieved March 9, 2019 from https://www.esri.com/library/whitepapers/pdfs/evaluating-population.pdf

Eurostat. *Demography and migration, Czech Republic*. Eurostat Database. Retrieved January 23, 2019, from https://ec.europa.eu/eurostat/web/population-demography-migration-projections/data/database

Eurostat (2015). *Population projection 2015, Czech Republic*. Eurostat Database. Retrieved August 9, 2018, from https://ec.europa.eu/eurostat/data/database

Eurostat Commission (2017). *Summary methodology of the 2015-based population projections.* Eurostat. Luxembourg. Retrieved March 20, 2019 from https://ec.europa.eu/eurostat/cache/metadata/Annexes/proj_esms_an1.pdf

Eurostat (2015). *Population projections 2015.* Retrieved February 21 2019, from https://ec.europa.eu/eurostat/web/population-demography-migration-projections/population-projections-data

Keyfitz N. (1982). *Can Knowledge Improve Projections?* Population and Development Review. *8*(4):729-751. doi:10.2307/1972470. Retrieved February 17, 2019 from https://www.jstor.org/stable/1972470?pq-origsite=summon&seq=2#metadata_info_tab_contents

Křesťanová J., Kurkin R. & Šafusová M. (2018). *Population Development in The Czech Republic in 2017.* Demografie 2018. *60*:313–332. Publications of Czech Statistical Office. Retrieved March 18, 2019 from https://www.czso.cz/documents/10180/61449036/13005318q4_313-332.pdf/086bbb07-0584-49a1-9a2a-61a2362f2ad5?version=1.0

Lee. R. D. (2012). *Projectioning Births in Post-Transition Populations: Stochastic Renewal with Serially Correlated Fertility.* American Statistical Association. (192.168.82.209). Retrieved

January 29, 2019, from https://pdfs.semanticscholar.org/7608/c9f9bfacc6afc003f6ef42b7e02f2fe678db.pdf

McDonald P. & Kippen R. (2008). *Projectioning Births Using a Three-Parameter Model.* The Australian National University. University of Melbourne. Retrieved January 28, 2019, from https://www.humanfertility.org/Docs/Symposium/McDonald_Kippen.pdf

Miskolczi M. & Cséfalvaiová K.(2013). *Process of Population Ageing and its Dynamic.* University of Economics, Prague. Retrieved March 18, 2019 from https://msed.vse.cz/files/2013/121-Miskolczi-Martina-paper.pdf

National Research Council (2000). *Beyond Six Billion: Projectioning the World's Population.* Washington, DC: The National Academies Press. 10.17226/9828. Retrieved February 15, 2019 from https://www.nap.edu/read/9828/chapter/1

Organization for Economic Co-operation and Development (2015). *International Migration Outlook 2015.* OECD Publishing, Paris. Retrieved January 29, 2019, from https://books.google.cz/books?id=6GuXCgAAQBAJ&printsec=frontcover&hl=ru#v=onepage&q&f=false

O'Neill B. C., Balk D., Brickman M. & Ezra M. (2001). *A Guide to Global Population Projections.* Demographic Research. *8*(4), 203-288. Retrieved February 26, 2019 from https://www.demographic-research.org/volumes/vol4/8/4-8.pdf

Pflaumer P. (1992). *Evaluating the Accuracy of Population Projections.* The 3rd International Conference on Social Science Methodology. Italy, 22 – 26 June 1992. Retrieved March 10, 2019 from https://www.researchgate.net/publication/294875252_Evaluating_the_Accuracy_of_Population_Projections

Planning Tank (2017). *Population projection | Meaning, Importance and Need.* Planning Techniques. Retrieved February 26, 2019 from https://planningtank.com/planning-techniques/population-projection

Population Reference Bureau (2001). *Understanding and Using Population Projections.* Population change. Retrieved February 15, 2019 from https://www.prb.org/understandingandusingpopulationprojections/

Renkou Y. (1980). *Some fundamental problems in population projection.* National Center for Biotechnology Information. (2):23-30. Retrieved February 13, 2019 from https://www.ncbi.nlm.nih.gov/pubmed/12159357

Smith K. & Shahidullah M. (1995). *An Evaluation of Population Projection Errors for Census Tracts.* Journal of the American Statistical Association. *90*(429):64-71. Retrieved March 10, 2019 from https://www.researchgate.net/publication/11228436_An_Evaluation_of_Population_Projection_Errors_for_Census_Tracts

Smith K. (2011). *We are seven billion.* Nature Climate Change. 1:331–335. Retrieved February 12, 2019, from https://www.nature.com/articles/nclimate1235

Swanson D.A., Tayman J. & Bryan T.M. (2001). *MAPE-R: a rescaled measure of accuracy for cross-sectional subnational population projections*. Journal of Population Research. *28*(225). Retrieved March 10, 2019 from https://link.springer.com/article/10.1007/s12546-011-9054-5#citeas

Šídlo L. & Tesárková K. (2009). *Vybrané možnosti hodnocení populačních prognóz. Aktuální populační prognózy České republiky*. Demografie 2. *51*(2):77-152. Retrieved August 15, 2018, from https://www.czso.cz/documents/10180/20565939/180309q2.pdf/4e9dbfae-e44c-47b1-ac88-c6faea58ee10?version=1.0

Šmejkalová L. (2011). *Hodnocení populační prognózy města Plzně.* Univerzita Karlova v Praze.

Tayman J., Smith S. K. & Rayer S. (2010). *Evaluating Population Projection Accuracy: A Regression Approach Using County Data.* Population research and policy review. *30*(2):235-262. Retrieved February 26, 2019 from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3061008/

World Population Review. *Czech Republic Population 2019*. Retrieved January 24, 2019, from http://worldpopulationreview.com/countries/czech-republic-population/

United Nations. *Population projection 2015, Czech Republic.* World population prospects. Retrieved October 11, 2018, from https://population.un.org/wpp/Download/Standard/Population/

Zhao F. (2010). *Evaluation of UN population projections and effect of urbanization on projection accuracy.* (763609862). Retrieved October 12, 2018 from https://search.proquest.com/docview/763609862?accountid=17203

Webster (2011). *Projectioning Population.* Retrieved March 6, 2019 from http://www.uobabylon.edu.iq/uobcoleges/ad_downloads/4_29239_397.pdf

Wilson T., Brokensha H., Rowe F. & Simpson, L. (2018). *Insights from the evaluation of past local area population projections.* Population Research and Policy Review. *37*(1):137-155. Retrieved February 13, from http://dx.doi.org/10.1007/s11113-017-9450-4

World Population Prospects (2015). *Methodology of the United Nations Population Estimates and Projections.* United Nations. New York. Retrieved March 20, 2019 from https://population.un.org/wpp/Publications/Files/WPP2015_Methodology.pdf