Charles University
Faculty of Science
Department of Demography and Geodemography

Study programme: Demography
Branch of study: Demography

Guldana Uteulina

Reproduction of the population of Kazakhstan in the data of the UN DESA Population Division and analysis of its forecasting estimate results from the years 1994-2019

Reprodukce obyvatelstva Kazachstánu v datech Populačního oddělení OSN a analýza jeho prognostických odhadů z let 1994-2019

Diploma thesis

Supervisor: RNDr. Tomáš Kučera, CSc.

Prague, 2024

Declaration

I declare that I have prepared the Thesis independently and that I have listed all the information sources and literature. Neither this Thesis nor any substantial part has been submitted for another or the same academic degree.

December 12, 2024, Prague …………………………

Acknowledgments

I would like to express my deepest gratitude to my thesis supervisor, RNDr. Tomáš Kučera, CSc., for his patience throughout my studies. I am immensely thankful for the countless opportunities he has provided and for the important life lessons he has shared, which I will carry with me well beyond my academic journey

Reproduction of the population of Kazakhstan in the data of the UN DESA Population Division and analysis of its forecasting estimate results from the years 1994-2019

Abstract

This thesis provides an analysis of the population of Kazakhstan and its reproduction characteristics since 1950 using United Nations World Population Prospects (UN WPP) data from 1994—2019. This analysis includes a comparison of estimates of population size, age—sex structure, and selected fertility, mortality and migration indicators, an analysis of differences in population figures across various datasets, and a comparative examination of UN WPP projections (medium fertility scenario) since 1994 against historical UN WPP estimates and the official demographic statistical data of Kazakhstan. Based on careful analysis, this study identifies differences between projected and actual population figures, revealing both overestimations and underestimations at different times. By tracing demographic trends in Kazakhstan from 1950 onwards and estimating the accuracy of the UN WPP medium fertility scenario projections, this thesis aims to emphasize the importance of context-specific data in demographic forecasting.

Keywords: population reproduction, population development, Kazakhstan, UN data, historical estimates, prognostic results, differences

Reprodukce obyvatelstva Kazachstánu v datech Populačního oddělení OSN a analýza jeho prognostických odhadů z let 1994-2019

Abstrakt

Tato diplomová práce poskytuje analýzu populace Kazachstánu a jejích reprodukčních charakteristik od roku 1950 s využitím dat Světových populačních prognóz OSN (OSN WPP) z let 1994–2019. Analýza zahrnuje srovnání odhadů velikosti populace, věkově-pohlavní struktury a vybraných ukazatelů plodnosti, úmrtnosti a migrace. Součástí práce je také analýza rozdílů v populačních údajích v různých datových souborech a komparativní zkoumání projekcí OSN WPP (scénář střední plodnosti) od roku 1994 v porovnání s historickými odhady UN WPP a oficiálními demografickými statistickými údaji Kazachstánu. Na základě důkladné analýzy studie identifikuje rozdíly mezi prognózovanými a skutečnými údaji o populaci, přičemž odhaluje jak nadhodnocení, tak podhodnocení v různých časových obdobích. Sledováním demografických trendů v Kazachstánu od roku 1950 a hodnocením přesnosti projekcí scénáře střední plodnosti OSN WPP si tato práce klade za cíl zdůraznit význam kontextově specifických dat v demografickém prognózování.

Klíčová slova: reprodukce obyvatelstva, populační vývoj, Kazachstán, data OSN, historické odhady, prognostické výsledky, rozdíly

CONTENTS

LIST OF ABBREVIATIONS

ASFR	Age Specific Fertility Rate
DHS	Demographic and Health Surveys
GNP	Gross National Product
LE	Life Expectancy
LSM	Living Standards Measurement Survey
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
ME	Mean Error
MPE	Mean Percentage Error
PE	Percentage Error
RMSE	Root Mean Square Error
RMSPE	Root Mean Square Percentage Error
TFR	Total Fertility Rate
UN	United Nations
UN DESA	United Nations Department of Economic and Social Affairs
USSR	The Union of Soviet Socialist Republics
WPP	World Population Prospects

LIST OF TABLES

LIST OF FIGURES

Chapter 1

Introduction

1.1 Problem definition

The United Nations Department of Economic and Social Affairs (UN DESA) presents population estimates and forecasts for countries worldwide through its World Population Prospects (WPP). However, the accuracy and reliability of these forecasts can differ, influenced by several factors including changes in fertility rates, migration, and mortality rates.

In the case of Kazakhstan, understanding the accuracy of these demographic forecasts is crucial due to its socio-economic transformations and demographic shifts since the dissolution of the Soviet Union. The UN DESA's population forecasts and historical estimates for Kazakhstan, covering the period from 1994 to 2019, present an opportunity to evaluate these forecasts' precision and their alignment with actual demographic outcomes as recorded in national statistics.

The primary problem this thesis addresses is the evaluation of the reliability of the UN DESA's demographic forecasts for Kazakhstan, identifying potential differences in population numbers and age-sex structures compared to Kazakhstan's official statistics. This involves dissecting the reasons behind any differences, which may stem from data collection methodologies, demographic model assumptions, or unforeseen demographic changes. Understanding these discrepancies is vital for improving future forecasts and refining demographic models to better support policy and planning processes in Kazakhstan and similar contexts.

1.2 Research goals and objectives

The primary research goal is to comprehensively analyze the demographic dynamics of Kazakhstan as presented in the UN DESA's World Population Prospects data, focusing on the accuracy and reliability of its forecasts from 1994 to 2019 revisions. This includes a comparison with historical data and official demographic statistics to assess the implications of population reproduction patterns.

The main research objectives include:

- Comparison of the historical estimates of Kazakhstan's population numbers and age-sex structure provided by WPP data with earlier records since 1950. This objective aims to trace the evolution of demographic characteristics over nearly seven decades to provide a context for more recent trends.

- Analyze the differences in population estimates between the UN's WPP data and official demographic statistics of Kazakhstan. This involves examining discrepancies in data collection methods, assumptions used in population modelling, and other factors that may influence population estimates.
- Comparison of the forecast results of the UN (using the medium fertility scenario) since 1994 with both the historical estimates and the official demographic statistics. This will help in evaluating the accuracy and reliability of UN projections and their utility for policy planning and demographic analysis.
- Assess the implications of identified differences and forecast errors on understanding the reproductive patterns and demographic changes in Kazakhstan.

1.3 Relevance of research

Demographic processes throughout history represent stages in a long-term population development trend, reflecting societal conditions and socio-economic progress. Central to sustainable societal development, human capital is influenced by population growth trends; either excessive growth or a significant decline can hinder development. Despite its vast area, Kazakhstan is among the smaller nations globally in population size, ranking 63rd worldwide and third among CIS countries, with a low population density of 7.3 people per km² as of early 2023.

To enhance its demographic profile, Kazakhstan implemented several state programs aimed at improving public health and managing migration effectively. Notably, the establishment of the Agency for Migration and Demography in 1997, which later merged its functions with other state bodies, helped steer demographic policies. Since 2003, these policies have fostered a steady increase in birth rates and natural population growth, helping to avert a demographic crisis.

National security laws in Kazakhstan identify demographic and public health deterioration, including declining birth rates and uncontrolled migration, as primary threats. The demographic strategy is thus a priority, supporting national security, especially given Kazakhstan's strategic Eurasian location.

The government actively addresses demographic challenges through the National Commission on Women's Affairs, Family and Demographic Policy. The global demographic imbalance is recognized as a critical issue, with strategies like "Kazakhstan - 2050" aiming to guide the nation's future development.

Despite these efforts, challenges such as potential aging due to increased retirement ages and continued emigration, particularly of European ethnic groups, remain. From 2003 to 2014, Kazakhstan saw positive migration balances due to socio-economic growth and effective policy implementation, although this trend reversed in later years due to economic factors and globalization.

Currently, Kazakhstan lacks a comprehensive, interdisciplinary approach to demographic research, with no official body dedicated to demographic forecasting.

Despite the critical role of demographic forecasting in planning and policy-making, significant discrepancies often exist between forecasted demographic data and actual outcomes.

The UN DESA provides population estimates and forecasts for countries worldwide through its World Population Prospects. However, the accuracy and reliability of these forecasts can vary, influenced by several factors including changes in fertility rates, migration, and mortality rates.

In the case of Kazakhstan, understanding the accuracy of these demographic forecasts is crucial due to its socio-economic transformations and demographic shifts since the dissolution of the Soviet Union. The UN DESA's population forecasts and historical estimates for Kazakhstan, covering the 1994 to 2019 revisions, present an opportunity to evaluate these forecasts' precision and their alignment with actual demographic outcomes as recorded in national statistics.

Chapter 2

Theoretical framework

2.1 Overview of literature

Population developments are inherently unpredictable. Yet we can make a forecast: a plausible and realistic assumption of future events based on existing information. However, the reliability and durability of such forecasts are limited. Since all forecasts have some degree of uncertainty, it is important to know more about this uncertainty (Keilman 1986).

Understanding the uncertainty in the forecasts is critical for decisions being made today, including infrastructure, investments in public and private pension funds, funding social insurance systems, and investments in mitigating and adapting to climate change (Christensen 2018).

The process of population forecasting is made up of several steps. It is worthwhile trying to find out in which stages of this process forecasting errors are usually generated. This can be done by distinguishing the following five types of errors (Keilman 1986):

1. Registration, rounding-off and estimation errors of the observed trends. In countries with poor data quality, such errors have a bearing on the quality of the population forecast.
2. Randomness in parameters. Stochastic fluctuations in the estimated numbers of births, deaths and migrants are not taken into account in the deterministic forecasting methods. However, a number of different studies have shown that errors caused by such simple extrapolations of (supposed) trends are very small (e.g. Sykes, 1969; Schweder, 1971).
3. Erroneous trends in forecasting parameters. If the quantitative course of real demographic trends differs substantially from the expected course, significant forecasting errors may result.
4. Sudden shifts in the parameters. Large forecasting errors can also be the result of extreme circumstances such as wars, disasters, serious economic depressions and the like. Such circumstances are not very common, however.
5. Inaccurate model specification. If one simulates the population trends over a known period with the aid of a forecasting model, the results should ideally correspond to the observations. In the past, the international migration component was not always incorporated in national population forecasts, often resulting in rather serious forecasting

errors. Now that international migration is taken into account, the models seem sufficiently valid.

The most profound forecasting errors are therefore caused by erroneous assumptions regarding trends in fertility, mortality and migration. Two stages in the assumption-making process account for this in particular. Firstly, when formulating future social and socio-demographic developments. For example, statements such as "the position of the family in society will become less dominant", or "structural factors such as advancing emancipation and growing individualism contribute to the continuing fertility decline; an increasing number of women are following higher education and wish to practice in their field for as long as they possible can. This trend is likely to continue". Secondly, important errors are generated when formulating the so-called key as assumptions, i.e. when social and socio-demographic hypotheses are translated into quantitative demographic indicators such as the average number of children per woman, the average life expectancy etc. It is clear that the hypotheses used in a population forecast, in particular general assumptions and key assumptions, have a greater bearing on the quality of the forecast than the accuracy of the model. If the key assumptions are correct, the model is of secondary importance. But if they do not adequately reflect the future, a good model cannot save the forecast.

Nathan Keyfitz assessed various established and rudimentary demographic theories: demographic transition, effects of development, Caldwell’s theory concerning education and fertility, urbanization, income distribution, Malthus’ writings on population, human capital, the Easterlin effect, opportunity costs, prosperity and fertility, and childbearing intentions. He tried to discover whether these theories had improved demographic forecasting, but his conclusion was negative. Although many of the theories are extensively tested, they have limited predictive validity in space and time, are strongly conditional, or cannot be applied without the difficult prediction of non-demographic factors (Keyfitz 1982 in Keilman 2019).

Significant uncertainty in projections arises from unforeseen events. Conflicts, natural disasters, economic downturns can lead to sudden influxes of migrants, temporarily reduce fertility rates, and cause numerous premature deaths. On the other hand, unexpected biomedical advancements can substantially increase life expectancy or offer new fertility solutions (NRC 2000).

In their article, UN DESA examines the sources of uncertainty in population projections surrounding the United Nations' global population estimates up to the year 2100 (United Nations 2019):

1. Future population trends are inherently uncertain, particularly over extended periods. While the future cannot be predicted with complete certainty, both recent and historical experiences can help inform assessments of likely outcomes in the short and long term. The uncertainty in population projections is largely due to the range of possible future trends in the three demographic components of population change: fertility, mortality, and international migration. In WPP, likely fertility and mortality trajectories have been evaluated for each country or region using demographic and statistical methods. Long-term population size is more uncertain because it depends not only on the

current population and the children they will have but also on the future fertility levels of their descendants.

2. The uncertainty regarding the future global population size is primarily driven by the unpredictability of the projected number of births. While the prediction interval for births widens considerably as the projection length increases, the prediction interval for global deaths remains relatively constant throughout the projection period.

A frequently overlooked source of uncertainty in population projections is the uncertainty surrounding current conditions, specifically the population size and structure, as well as fertility and mortality rates in a given country at the start of the projections. While this data is usually well-documented in industrialized nations, significant information gaps exist, particularly in Africa and parts of Asia. For many countries, data is derived only from sample surveys, and in some cases, even such information is unavailable. Although the UN Population Division strives to produce population estimates for all countries, the need to adjust figures from the recent past with each new assessment as additional data becomes available illustrates the complexity of this task. Despite this, demographers have become accustomed to having precise point estimates for demographic indicators, even in countries with very limited empirical data. Consequently, the inherent uncertainty of these indicators concerning the initial conditions is often overlooked when making projections (Lutz 2010).

Error in projecting a country's total population is generally accompanied by errors in regard to the sizes of particular age groups. Projections of the youngest and the oldest age groups tend to be the least reliable. In the past, these errors have been the result of too high projections of fertility (resulting in too many infants and young children) and too high projections of mortality (resulting in too few elderly) (NRC 2000).

Forecast accuracy is just one aspect of evaluating forecasts. Smith and Swanson (2001) identify several criteria for assessing forecasts, including the provision of necessary detail, face validity, plausibility of production, timeliness, ease of application and explanation, political acceptability, and forecast accuracy. Although forecast accuracy is arguably the most critical criterion (Wilson 2007), it is essential to consider the broader context of these assessments.

As demographic data for forecast years becomes available, it allows for evaluating the accuracy of earlier forecasts (Wilson 2007). Such evaluations can be beneficial in various ways: they may identify systematic errors that forecasters could address in future predictions; they provide guidance on the likely accuracy of current population forecasts, assuming that current forecasting efforts are comparable to past efforts (Keyfitz 1981); and they can directly inform or calibrate models used to generate predictive intervals for probabilistic population forecasts (Keilman, et al. 2001).

A number of research evaluates national population forecasts by comparing them retrospectively with actual statistics. These analyses, conducted by various scholars, indicate that forecasts tend to be more accurate over shorter durations and for larger populations. However, they also reveal significant variations in forecast accuracy across different regions and components (Keilman 2008). Due to the extrapolative nature of population forecasting, anomalies in the data often result in substantial errors. Additionally, data quality is closely linked to forecast accuracy, with poorer data leading to less reliable predictions.

When projections are constantly examined, analyzing accuracy may seem ineffective. However, unless an agency explicitly acknowledges that its assumptions are unrealistic, most users will interpret projection results as forecasts, viewing them as probable future developments based on current understanding, as noted by Keyfitz in 1972.

Since 2015, the U.N. has incorporated probabilistic projections of the total fertility rate into their official population forecasts for all countries, using the Bayesian hierarchical model. Despite this advancement, a significant source of uncertainty remains unaddressed: uncertainty regarding historical fertility levels (Liu and Raftery 2018). This is a minor concern for countries with well-established vital registration systems, which applies to fewer than half of the approximately 200 countries worldwide. For the majority without such systems, this uncertainty can be substantial.

While the new U.N. method more systematically addresses uncertainty than previous approaches, it still overlooks certain sources of uncertainty. The Bayesian hierarchical model relies on estimates of current and historical populations and fertility and mortality rates. In countries with high-quality vital registration systems that provide accurate birth and death counts, such uncertainty is minimal. This applies to approximately 80 out of the world's 200 or so countries. However, the other 120 countries rely on fertility and mortality estimates derived from surveys, such as those conducted by the DHS. Although DHS data is one of the most reliable sources, it has been criticized for significantly underestimating the TFR in some Sub-Saharan African countries (Liu and Raftery 2018).

Chapter 3

Data availability

3.1 Data sources

This study spans the timeframe from the 1994 revision through to the 2019 revision (13 editions), utilizing the UN WPP, produced by the Department of Economic and Social Affairs Population Division. These biennial publications provide population estimates and projections for all countries worldwide, with the first being released in the early 1950s.

The UN WPP has become an integral tool within the entire United Nations framework and is widely utilized by various international organizations and academic researchers. Its role is crucial in development planning, monitoring, and global modeling efforts. In the context of Kazakhstan, where there is no established institution systematically producing national forecasts, the significance of these global population estimates becomes even more pronounced. These projections are vital for the country's economic and social planning for the coming decades. Local data for this study was sourced from Kazakhstan's National Bureau of statistics, providing a foundational base for comparison and analysis.

Tab. 1: Data Sources

Revision	Total population	Total fertility rate	Infant mortality	Life expectancy at birth	International migration
1994	(1990) Based on official estimate consistent with the 1989 census.	Based on registered births, by age of mother, to 1990.	Based on registered births and infants deaths to 1990, adjusted upward by 25 per cent due to omitted infant deaths.	Estimated from registered deaths adjusted for underregistration and the underlying populations, both by age and sex, to 1990.	(1990-1995): Based on estimates of international migration to 1990, including migration between the republics of the former USSR, and on assumed future trends.
1996	(1995) Estimated to be consistent with the 1989 census.	Based on registered births, by age of	Based on registered births and infants' deaths	Estimated from registered deaths adjusted for underregistration	(1990-1995): Based on estimates of net migration

		mother, through 1990.	through 1993, adjusted upward by 25 per cent due to omitted infant deaths prior to 1991.	and the underlying populations, both by age and sex, through 1991.	through 1993 and assumed subsequent trends.
1998	Estimated to be consistent with the 1989 census.	Based on official estimates of total fertility available through 1998. Estimates from the 1999 Kazakhstan DHS were also considered.	Based on registered births and infants deaths through 1993, adjusted upward by 25 per cent due to omitted infant deaths prior to 1991.	Based on an official life table for 1997.	Based on estimates of net migration through 1993 and assumed subsequent trends.
2000	Estimated to be consistent with the 1989 census.	Based on official estimates of total fertility available through 1998. Estimates from the 1999 Kazakhstan DHS were also considered.	Based on births and infant deaths registered through 1998, adjusted upward by a factor of 12 before 1985, decreasing to 1.1 for 1990-1993, to take account of infant deaths omitted owing to the use of a definition of infant death that does not conform to international standards. Estimates from 1999 Kazakhstan DHS were also considered.	Based on an official life table for 1997.	Based on estimates of net international migration derived as the difference between overall population growth and natural increase through 1999.
2002	Total population (2000): Estimated to be consistent with the 1999 census and with estimates of the subsequent trends in fertility, mortality and international migration.	Based on official estimates of total fertility available through 1998. Estimates from the 1999 Kazakhstan DHS were also considered.	Based on births and infant deaths registered through 1998, adjusted upward by a factor of 1,2 before 1985, decreasing to 1,1 for 1990-1993, to include infant deaths omitted owing to the use of a	Based on an official life table for 1997.	Based on estimates of net international migration derived as the difference between overall population growth and natural increase through 1999.

			definition of infant death that does not conform to international standards. Estimates from 1999 Kazakhstan DHS were also considered.		
2004	Total population (2005): Estimated to be consistent with the 1999 census and with estimates of the subsequent trends in fertility, mortality and international migration	Based on the 1999 Kazakhstan DHS and on official total fertility estimates available through 2002.	Infant and/or child mortality: Based on maternity history data from the 1999 Kazakhstan DHS.	Life expectancy at birth: Based on a life table calculated from registered deaths by age and sex for 2003, adjusting infant and child mortality rates to be consistent with rates from the 1999 Kazakhstan DHS.	International migration: Based on estimates of net international migration, derived as the difference between overall population growth and natural increase through 2003.
2006	Total population (2005): Estimated to be consistent with the 1999 census, with an official estimate for 1 July 2005 and with estimates of trends in fertility, mortality and international migration.	Based on the 1999 Kazakhstan DHS and official estimates of total fertility available through 2004.	Infant and/or child mortality: Based on data on children ever born and surviving classified by age of mother from the 2006 Kazakhstan MICS.	Life expectancy at birth: Based on a life table derived from registered deaths by age and sex for 2003, adjusting infant and child mortality rates to be consistent with rates from the 2006 Kazakhstan MICS.	Based on estimates of net international migration derived as the difference between overall population growth and natural increase through 2005.
2008	Estimated to be consistent with the 1999 census and with estimates of the subsequent trends in fertility, mortality and international migration	1995, 1999 DHS	Infant and/or child mortality: Based on data on children ever born and surviving classified by age of mother from the 2006 Kazakhstan MICS.	Life expectancy at birth: Based on a life table derived from registered deaths by age and sex for 2003, adjusting infant and child mortality rates to be consistent with rates from the 2006 Kazakhstan MICS.	Based on estimates of net international migration derived as the difference between overall population growth and natural increase through 2008.
2010	Total population (2009): Estimated to be consistent with the 1959, 1970, 1979, 1989, 1999 and 2009 census, and with estimates of the	Based on: (a) official estimates of total fertility available through 2008, (b) maternity-history data	Based on: (a) births and infant deaths registered through 2008 adjusted to compensate for infant deaths omitted	Based on official estimates of life expectancy available through 2008 adjusted for underreporting of infant and child mortality. The age pattern of	Based on estimates of net international migration derived as the difference between overall

	subsequent trends in fertility, mortality and international migration.	from the 1995 and 1999 Kazakhstan DHS,	owing to the use of a definition of infant death that does not conform to international standards, (b) data on children ever born and surviving classified by age of mother from the 1989 and 1999 censuses, 1996 LSMS and 2006 MICS, (c) maternity-history data from the 1995 and 1999 DHS.	mortality is derived from a life table constructed on the basis of 2005 data.	population growth and natural increase through 2008.
2012	Estimated to be consistent with the 1959, 1970, 1979, 1989, 1999 and 2009 census, and with estimates of the subsequent trends in fertility, mortality and international migration.	Based on: (a) official estimates of total fertility available through 2009, (b) maternity-history data from the 1995 and 1999 Kazakhstan DHS, and (c) births in the preceding 12 months to the 2010 MICS classified by age of mother	Based on official estimates of life expectancy available through 2008 adjusted to take into account underreporting of infant and child mortality. The age pattern of mortality is derived from a life table constructed on the basis of 2005 data.	Based on: (a) data on children ever born and surviving classified by age of mother from the 1989 and 1999 censuses, 1996 LSMS and 2006 MICS, (b) maternity-history data from the 1995 and 1999 DHS, and (c) based on data on children ever born and surviving from the 2010 MICS. Official estimates of mortality through 2010 were also considered.	Based on estimates of net international migration derived as the difference between overall population growth and natural increase through 2010.
2015	Estimated to be consistent with the 1959, 1970, 1979, 1989, 1999 and 2009 censuses, and with estimates of the subsequent trends in fertility, mortality and international migration.	Based on: (a) official estimates of total fertility available through 2011, (b) maternity-history data from the 1995 and 1999 Kazakhstan DHS, and (c) births in the preceding 12 months from	Based on official estimates of life expectancy available through 2008 adjusted for underreporting of infant and child mortality. The age pattern of mortality is derived from a life table based on	Based on: (a) data on children ever born and surviving classified by age of mother from the 1989 and 1999 censuses, 1996 LSMS and 2006 MICS, (b) maternity-history data from the 1995 and 1999 DHS, (c) data on children ever born and surviving from	Based on estimates of net international migration derived as the difference between overall population growth and natural increase through 2014.

		the 2010 MICS classified by age of mother	2010-2013 data.	the 2010 MICS, and (d) official estimates of mortality through 2013.	
2017	Estimated to be consistent with the 1959, 1970, 1979, 1989, 1999 and 2009 censuses, with official population estimates through 2014, and with estimates of the subsequent trends in fertility, mortality and international migration.	Based on: (a) official estimates of total fertility available through 2015, (b) maternity-history data from the 1995 and 1999 Kazakhstan DHS, and (c) births in the preceding 12 months from the 2010 and 2015 MICS classified by age of mother.	Infant and child mortality: Based on: (a) data on children ever born and surviving classified by age of mother from the 1989 and 1999 censuses, 1996 LSMS and 2006, 2010 and 2015 MICS, (b) maternity-history data from the 1995 and 1999 DHS, and (c) official estimates of mortality through 2013.	Life expectancy at birth: Based on official estimates of life expectancy available through 2012 adjusted for underreporting of infant and child mortality. The age pattern of mortality is derived from a life table based on 2010-2013 data.	Based on estimates of net international migration derived as the difference between overall population growth and natural increase through 2014.
2019	Total population and distribution by age and sex estimated to be consistent with the population by age and sex of the (a) 1959, 1970, 1979, 1989, 1999, 2009 censuses; (b) official estimates through 2017; and with estimates of the subsequent trends in fertility, mortality, and international migration.	Total fertility rate and age pattern of fertility based on: (a) official estimates of age-specific fertility rates through 2017; (b) registered births classified by age of mother and the underlying female population by age through 2017; (c) birth-histories data from the 1995, 1999 DHS; (d) births in the household in the preceding 12 (or 24) months	Child Mortality: Under-five mortality based on: (a) official estimates of infant and child mortality through 2017; (b) registered infant and child deaths through 2017; (c) direct estimates from births and deaths under-five calculated from full birth histories data from the 1995, 1999 DHS; (d) indirect estimates from data on children ever born and surviving from the 1989, 1999 censuses and 1996 LSMS and 2006, 2010-	Overall Mortality: Life expectancy at birth and age pattern of mortality based on: (a) official estimates through 2017; (b) registered deaths by age and sex available through 2017; (c) adjusted for underreporting of infant and child deaths	International migration based on: estimates derived as the differences between overall population growth and natural increase.

		classified by age of mother from the 2010-2011, 2015 MICS.	2011, 2015 MICS.		

Source: UN WPP 1994–2019 revisions

3.2 Data quality

Reliable and precise foundational data on demographic elements such as population figures, age distribution, fertility rates, mortality rates, and net migration are essential for the creation of dependable demographic forecasts, as highlighted by Keilman (1990). The quality assessment of data from the UN World Population Prospects, as previously noted, encounters certain challenges regarding its dependability. The initial revisions of the WPP, particularly from 1994 to 1996, which marked the emergence of newly independent states, including Kazakhstan, faced complex estimation issues due to only having limited data available at that time.

For the 1990 population figures, the estimates are aligned with the official numbers from the 1989 census. Fertility rates up to 1990 were based on officially registered births by the mother's age. Life expectancy calculations up to 1990 were derived from death registrations, corrected for underreporting and factoring in population composition by age and gender. International migration numbers up to 1990 were based on international migration estimates, including intra-USSR republic movements and projected future trends as per the 1994 revision of the UN WPP.

Later revisions, starting from 2002, incorporated more current data, such as the 1999 census for total population and various health surveys for other indicators like fertility rates. Mortality rates were extracted from the 1991 death registrations, with adjustments made for infant mortality based on revised figures. Net migration rates were deduced from the difference between actual population growth and the estimated natural increase.

Historical demographic estimates are often revised in later years to improve their accuracy. When projections are initially developed, data on a country's base population are derived from either the latest census or a comprehensive national population registry, representing the most accurate assessment available for the base year's age structure at that time. Additionally, projections for future fertility, mortality, and migration levels are created by extending historical trends of these vital statistics and incorporating a detailed understanding of how socio-economic factors, cultural dynamics, and policy interventions might influence these changes (Keilman 2001). The accuracy of these historical estimates can be enhanced with the acquisition of new data, such as recent census results or post-census evaluations, which may refine the initial figures. Improved methodologies for estimating demographic variables, along with more sophisticated analyses of fertility patterns, mortality rates, and migratory movements, further contribute to the refinement of past estimates.

Chapter 4

UN Population Division's Methodology

Each revision of the United Nations' official population estimates and projections involves a thorough process that comprises two primary steps: updating the demographic information for each country, which may entail revising past estimates with new data, and formulating detailed assumptions about future trends in fertility, mortality, and international migration for every country (United Nations 2019).

Since 1963 Revision, the Population Division has employed the cohort-component projection method for producing individual country projections. This method, the most common projection method used by demographers, provides an accounting framework for the three demographic components of change — fertility, mortality and international migration — and applies it to the population in question. Technically, it is not a complete projection method on its own, as it requires that the components of change — fertility, mortality and migration — be projected in advance. 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 (Preston et al., 2001 in United Nations 2019).

The historical data is generated using the cohort-component method, which advances the population in five-year intervals based on age-specific demographic data. This method uses census data to anchor population counts by age and sex, with adjustments made for data deficiencies such as inaccurate age reporting or underreported vital events (United Nations 2019).

The jump-off year marks the transition from historical estimates to future projections. The UN Population Division employs various projection variants to illustrate potential impacts of different fertility scenarios on population size and structure. More recently, probabilistic methods have been introduced for certain demographic components like total fertility and life expectancy, providing a median trajectory along with prediction intervals to indicate uncertainty (United Nations 2017).

A crucial goal of each revision is to maintain the consistency and comparability of demographic data across time and countries. New data undergo careful quality checks and impact analyses to assess changes in fertility, mortality, and migration trends. The projection phase uses statistical techniques and guidelines to project future demographic changes, with occasional deviations for specific countries, especially in projecting net international migration and life expectancy (United Nations 2017).

4.1 The overall approach and key steps involved in generating population estimates

The UN World Population Prospects methodology involves a continuous process of updating and revising demographic estimates with each new edition (United Nations 2019). This "re-estimation" is prompted by the latest data from sources like censuses, demographic surveys, vital event registries, and other relevant information, including migration and refugee statistics. When new data are available, the time series for fertility, mortality, and migration, as well as trends in age and sex distribution, are extended and corrected as needed.

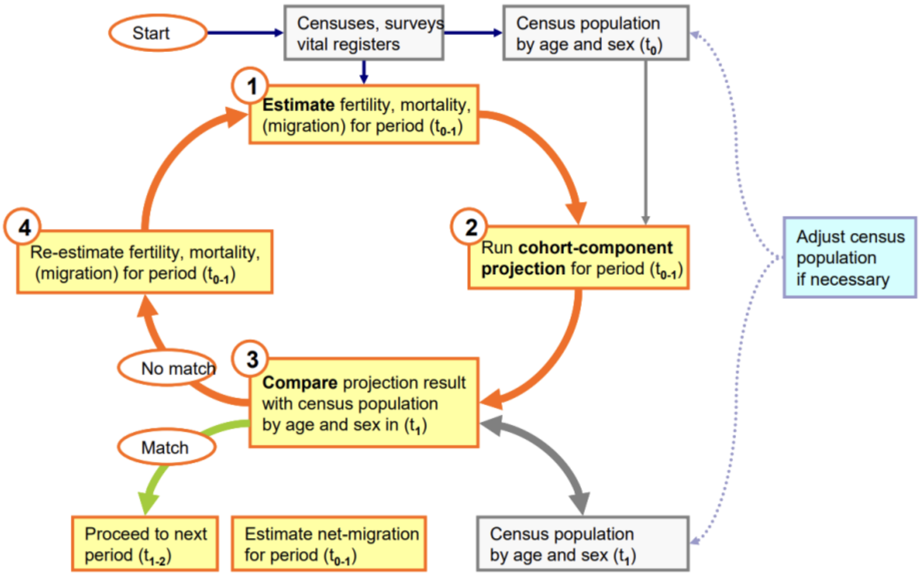
In more developed regions, the availability of detailed fertility and mortality data and regular censuses simplifies generating reliable past population dynamics (United Nations 2019). The basic demographic equation, which incorporates births, deaths, immigration, and emigration, allows for accurate intercensal and postcensal population estimates, provided that vital statistics and migration data are available and accurate. In cases where migration data are lacking, net international migration estimates are made by comparing census-based population growth with growth indicated by fertility and mortality rates.

In less developed regions, where demographic data are often limited or unreliable, indirect estimation methods based on models are used to derive more accurate figures (United Nations 2002). Revising demographic estimates for each country involves evaluating the latest data on fertility, mortality, and international migration, along with new census data that provide insights into the population age distribution (Gerland and others 2014).

The process of revising estimates involves not only assessing the quality of the available data but also ensuring consistency across various datasets to uphold the demographic balancing equation, which includes age and sex distributions as cohorts evolve over time. This approach aims to accurately reflect past trends in fertility, mortality, and migration in relation to changes in population size and structure (United Nations 2019).

The methodology used in the 2019 revision of the World Population Prospects, as well as in previous editions, follows a structured four-step process. This process, illustrated in the figure below, outlines how estimates for fertility, mortality, and net migration are rigorously tested for internal consistency within a cohort-component projection framework covering the period from t_0 to t_1. This systematic approach is a standard procedure in each new revision of the World Population Prospects. As new data become available, past demographic estimates are reassessed and potentially adjusted, ensuring that each revision reflects the most accurate and up-to-date understanding of past demographic trends (United Nations 2019).

Fig. 1: The method employed to maintain consistency between demographic components and the overall population during intercensal period.



Source: UN WPP 2019 revision.

4.2 The preparation of population projections

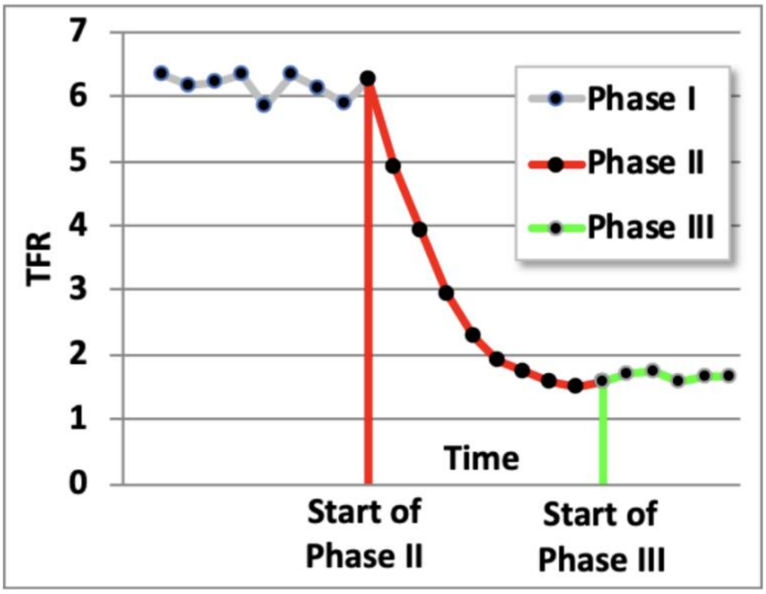
Since the 1963 Revision, the Population Division has employed the cohort-component projection method to create projections for individual countries (United Nations 2019). This method, widely adopted by demographers, provides a framework for managing the three main demographic components: fertility, mortality, and international migration. It applies these components to the population under study (United Nations 2019). Technically, this method is not a standalone projection tool, as it relies on prior projections of demographic changes. Instead, it utilizes matrix algebra to assist demographers in estimating how assumed future trends in fertility, mortality, and international migration will affect the population at a future date.

4.2.1 Fertility assumptions

The World Population Prospects consider multiple scenarios based on different fertility assumptions to account for potential demographic changes. These include a medium-fertility scenario, a high-fertility scenario, a low-fertility scenario, a constant-fertility scenario, an instant-replacement scenario, and a momentum variant that adjusts mortality and migration assumptions differently than the instant-replacement-fertility scenario. In this study, the analysis of UN projections focuses on the medium-fertility assumption to explore and understand potential population trends.

The 2019 revision of the World Population Prospects incorporated probabilistic techniques for forecasting total fertility, a method first introduced in the 2010 revision and further refined in subsequent updates (United Nations 2019). These methods analyze fertility levels and trends from 1950 to 2020 for all countries, using the demographic transition theory as a framework to predict future country-specific fertility rates. Historically, fertility evolution is understood to occur in three phases: an initial high-fertility phase, a transition phase where fertility rates decline, and a final low-fertility phase (United Nations 2019).

Fig. 2: Phases of fertility evolution.



Source: Alkema and others, 2011.

The transition to the second phase of the fertility transition is marked by the peak fertility rates. If the peak fertility rate was below 5.5 births per woman before 1950, the country is considered to have entered phase II prior to that year. Otherwise, the transition begins at the peak fertility period. Phase II concludes when fertility rates drop below 2 births per woman and remain there. By 2019, it was confirmed that all countries had either entered or completed their fertility transition.

The 2019 revision of the World Population Prospects focused on the latter two phases of the fertility transition, excluding the pre-transition phase. It observed that the decline in fertility rates tends to accelerate initially before slowing as it approaches replacement levels. This consistent pattern across countries enables demographers to more accurately predict the rate of fertility decline, enhancing the accuracy of future fertility projections.

Although the rate of fertility decline varies by country, a general trend has been identified worldwide. Typically, fertility rates decline rapidly at first, then slow down as they approach intermediate levels and the replacement threshold. Differences in this pattern, especially during the early and late stages of the transition, have been documented. These observations help demographers model the rate of fertility decline over specific periods, providing a strategic method for estimating future trends rather than simply predicting exact future rates (United Nations 2006).

The probabilistic method for projecting total fertility, first implemented in the 2010 revision, includes two main components (United Nations 2019). The first component models the transition from high to low fertility (phase II) using a double-logistic decline function, with speed and pattern estimated through a Bayesian hierarchical model (BHM). This approach generates country-specific predictions based on both the country's historical fertility trends and those of other countries that have experienced similar declines. This method captures both systematic declines and random fluctuations in fertility rates, offering a comprehensive view of potential future trends.

The second component is applied once a country reaches the low-fertility phase (phase III), using a time-series model to project future changes. This model assumes that fertility will stabilize at specific levels, informed by empirical data from countries that have experienced fertility rebounds from below-replacement levels. These projections are based on the global experience, particularly from European and East Asian countries, which have shown varying patterns of fertility recovery, providing insights into expected long-term fertility behavior for each country.

For countries that had not yet transitioned to phase III by the 2015–2020 period, the Bayesian hierarchical model created 186 000 double-logistic curves to represent the variations in the fertility decline function for countries undergoing a fertility reduction. This extensive set of curves allowed for the calculation of 100 000 total fertility projections, incorporating a distortion term to account for uncertainty in the decline model.

Previously, until the 2010 revision, projections assumed a long-term fertility rate of 1,85 children per woman. Recent revisions have included scenarios where fertility rates could drop below this figure, reflecting uncertainties about the historical minimum fertility levels before any potential recovery in phase III. The timing of the end of phase II and the beginning of phase III, as well as the rate and extent of fertility changes, differ among the projected trajectories. Countries that had entered phase III by 2015–2020 used a time-series model specific to that phase.

Looking ahead, projections include scenarios where countries are in either phase II or III, with the expectation that all will eventually reach phase III. By 2095-2100, the unweighted global average fertility rate, based on 40 countries in phase III, is projected to be 1,78 children per woman, with a 95% prediction interval ranging from 1,66 to 1,94 children per woman (United Nations 2019).

The fertility forecasts in the 2019 revision are based on historical fertility trends, assuming that the factors driving fertility decline will persist. However, these projections could be influenced by future changes. For example, if there is a significant expansion in access to family planning resources, the current median fertility estimates might be higher than the actual future rates. Conversely, a decrease in the use of modern contraceptives or a renewed preference for early marriage and larger families could lead to underestimating future fertility levels (United Nations 2019).

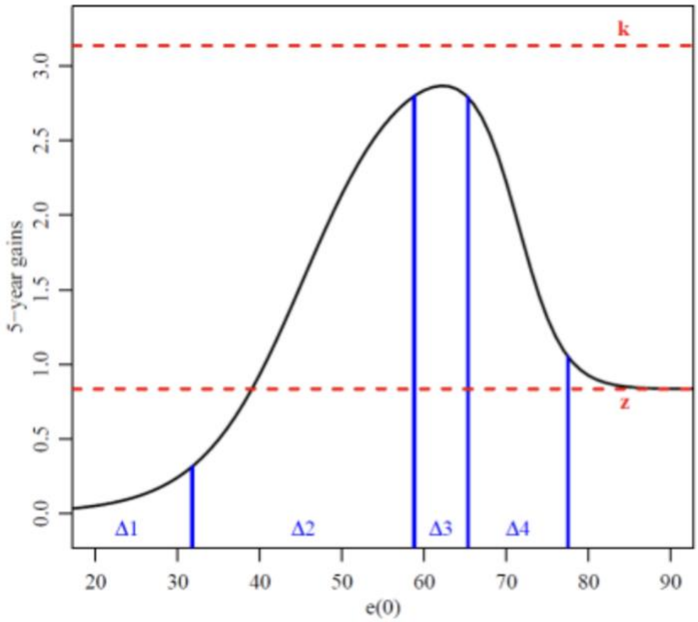
4.2.2 Mortality assumptions

The mortality projections in the World Population Prospects are based on assumptions related to life expectancy at birth, with distinctions made by sex (United Nations 2019). The Population Division provides 80% and 95% prediction intervals for future life expectancy levels, along with a median trajectory derived from a statistical model tracking changes in mortality over time. This median trajectory guides the mortality trend across various scenarios, including high, medium, and low fertility, instant-replacement-fertility, and zero-migration scenarios, all of which focus on fertility rate variations. Life expectancy is generally projected to increase throughout the forecast period.

The 2019 revision introduced probabilistic methods for projecting life expectancy at birth, building on model enhancements first introduced in the 2017 revision. These methods offer a more comprehensive and statistically robust approach to predicting future life expectancy trends.

In the 2019 revision, life expectancy projections utilize probabilistic methods through two distinct models. The first model represents the gradual increase in female life expectancy, depicting the transition from higher to lower mortality rates using a logistic function divided into two phases (Raftery et al. 2013).

Fig. 3: Phases of the mortality transition: gains in life expectancy at birth by level of life expectancy at birth (years)



Source: Raftery et al., 2013.

In the initial phase, life expectancy growth is slow, primarily due to improvements in hygiene and nutrition. This phase is followed by a rapid improvement phase driven largely by public health measures and medical care targeting infant and child mortality, such as better sanitation, infant feeding practices, and widespread immunization programs. Once the most accessible gains from combating infectious diseases in children are achieved, the subsequent phase begins. This phase continues to fight infectious diseases across all age groups while also addressing chronic non-communicable diseases prevalent in older populations. Because preventing deaths from non-communicable diseases is complex and the increase in life expectancy from saving older individuals is relatively smaller compared to children, the pace of life expectancy growth slows during this second phase. This model reflects the nuanced progress in life expectancy as societies advance in healthcare and disease prevention across different stages of life.

In countries undergoing a mortality transition, improvements in life expectancy at birth are influenced by two key components within the model:

1. *Systematic Gains*: The model uses a double-logistic improvement function to represent systematic gains in life expectancy. This function, refined in earlier editions of the World Population Prospects (United Nations 2006), models life expectancy improvements based on the current level of life expectancy. Parameters are derived from observed changes in female life

expectancy from 1950 to 2020, calculated using a Bayesian hierarchical model. This model provides country-specific distributions for these parameters and future life expectancy trends, combining specific country data with aggregated global data to enhance accuracy.

2. *Random Distortion*: For each country or region, future projections of life expectancy are calculated for subsequent five-year periods using a random walk with drift. This method accounts for random variations and incorporates a drift parameter to predict gradual changes over time based on historical data.

The second model used to forecast future mortality trends focuses on the difference between female and male life expectancy at birth. Male life expectancy projections are derived by adjusting female life expectancy figures based on the projected gender gap. This method considers the correlation between life expectancies of both genders and accounts for anomalies during periods of crisis or conflict. The gender gap is assessed using an autoregressive approach, with female life expectancy as a significant predictor. Numerous studies have explored the biological, behavioral, and socioeconomic factors contributing to this life expectancy gap between genders. Recent trends suggest a closing of this gap in almost all high-income nations and some emerging economies, likely due to effective public health initiatives and medical advancements (United Nations 2019).

The projection model suggests that the gender gap in life expectancy typically widens in countries with lower life expectancies but narrows as female life expectancy surpasses approximately 75 years. This trend is expected to continue until female life expectancy reaches around 86 years. At very high levels of female life expectancy (around or above 86 years), the gap is modeled to stabilize, incorporating standard deviations to account for uncertainties due to limited data on factors affecting changes in the gap at these higher ages. This approach leads to some alignment between male and female life expectancy projections within certain countries during the forecast period (United Nations 2019).

4.2.3 International migration assumptions

Projecting international migration is particularly challenging due to the unpredictable nature of migration and often incomplete historical data (United Nations 2019). Migration patterns are influenced by rapidly changing economic, social, political, and environmental factors, causing historical trends to shift dramatically within a short time. Some countries may transition from being primarily sources of migrants to becoming destinations, and vice versa.

Historically, migration has had a minor impact on the size and composition of national populations, allowing for the assumption that migration rates will remain constant for most of the projection period. However, in countries where migration is a significant demographic factor, a more nuanced approach is required (United Nations 2019).

International migration involves individuals leaving one country (emigration) and entering another (immigration). Ideally, migration is studied as a flow between countries, but comprehensive data on these flows are only available for a few nations. As a result, the 2019 revision, like previous ones, considers international migration as net migration. This metric represents the difference between the number of immigrants and emigrants over a specific period, reflecting the net effect on the population size and structure of both origin and destination countries. Even if large numbers of people immigrate to and emigrate from a country, net migration can be zero if the numbers balance out (United Nations 2019).

Several key sources of information inform assumptions about future international migration trends: data on net migration flows and their components (immigration and emigration), labor migration statistics, estimates of undocumented or irregular migration, movements of refugees and asylum-seekers, and changes in the foreign-born population (United Nations 2019).

For most countries, projecting future net international migration involves distinguishing between regular international migrants and refugees. It is generally assumed that recent stable migration levels will persist until around 2045–2050. Other factors include government migration policies and estimates of undocumented movements. For refugees, the common assumption is that they will return to their countries of origin within five to ten years.

Migration assumptions typically focus on the net number of migrants, considering gender distribution based on known participation in different migration types, such as labor migration or family reunification. Due to limited age distribution data, models allocate the total number of male and female migrants by age based on the predominant migration type. These age and sex profiles are then integrated into the cohort-component projection model (United Nations 2019).

For countries known for temporary labor migration, models account for the return of these migrants as they age. The same approach applies to refugee flows. For a few countries with minimal international migration, net migration is assumed to be zero soon after projections begin. However, most countries are expected to have non-zero net migration, with more projected as net senders than net receivers (United Nations 2019).

To ensure global consistency, the total of all net international migration flows is adjusted to zero for each five-year projection period. This involves an iterative process where estimates for individual countries, especially those with the highest net migration (both positive and negative), are continuously refined (United Nations 2019).

In the normal migration scenario, future net international migration projections are based on historical data and each country's policy attitudes towards migration. Typically, projected net migration levels remain consistent until 2045–2050, unless affected by significant recent changes in migration patterns, refugee movements, or temporary labor migrations. Post-2050, net migration is expected to continue at the 2045–2050 levels until 2095–2100. This approach serves as a practical middle ground, recognizing the difficulties in precisely forecasting long-term migration for every country while acknowledging that net migration is unlikely to cease entirely (United Nations 2019).

According to the 2019 WPP revision, international migration has become a significant factor influencing population dynamics in various regions. However, projecting future migration levels remains uncertain due to potential significant fluctuations driven by labor market demands, policy changes, conflicts, and environmental crises, which are difficult to predict.

Under the zero-migration assumption, net international migration for each country is set to zero starting in the 2015–2020 period, removing migration flows from future population projections. While this assumption simplifies projections, it does not accurately reflect the dynamic nature of migration patterns (United Nations 2019).

4.3 Methods for Population Projections

The Population Division employs the cohort-component method, established by the United Nations in 1956, to analyze and project the population dynamics of each country up to the year 2100 (United Nations 2019). This method is essential for maintaining internal consistency over time and across demographic components (fertility, mortality, and net migration), ensuring an accurate representation of population size and composition by age and sex. Using this approach, population projections are made from 2020 to 2100 in five-year intervals. These projections incorporate future survival probabilities, net migration, and the number of births, all segmented by age and sex. Each subsequent population figure for a five-year period is calculated by adding births and net migration to the previous period's population and subtracting deaths, with these calculations stratified by sex and age group.

The Population Division releases two types of population projections: probabilistic and deterministic. The probabilistic version adds a layer of complexity to the deterministic model by introducing uncertainty in future fertility and mortality changes. However, these probabilistic projections do not account for uncertainties in future net migration or baseline demographic rates.

In contrast, the deterministic version uses the median trajectory for fertility and life expectancy, along with previously discussed assumptions, to project population figures, adhering strictly to the expected paths without factoring in potential variability (United Nations 2019).

The table below chronologically organizes 13 revisions of the World Population Prospects, each marked by its publication year and the corresponding base year used for projections. Each subsequent revision follows this methodology, updating the base year for projecting future demographic shifts. When multiple revisions share the same base year, they are distinguished by labels such as I, II, III. For example, the revisions of 2008, 2010, and 2012 all use 2010 as their base year and are labeled as 2010I, 2010II, and 2010III respectively. Additionally, the table highlights any changes in methodology compared to the previous revisions.

Tab. 2: United Nations World Population Prospects analyzed in the study.

Year of Publication	Jump-off year	Label	Notes (Methodology change)
1994	1990	1990	-
1996	1995	1995I	The overall methodology used in the 1996 Revision remained consistent with that of the 1994 Revision.
1998	1995	1995II	The 1998 Revision incorporated fresh data and updated assumptions. Notably, it included new insights into current fertility rates and dropped the previous assumption that fertility in countries currently below replacement level would increase to replacement level by 2050.
2000	2000	2000I	Fertility: • High-fertility countries may not achieve replacement-level fertility by 2050. Migration: • For many countries, net international migration is projected to remain above zero throughout the fifty-year projection period. • A detailed analysis of international migration data was conducted to generate past flow estimates and evaluate future migration prospects at the country level. Two additional projection variants were introduced to explore the effects of sustained net international migration and decreasing mortality on population growth and aging.
2002	2000	2000II	Fertility:

			• In the medium variant, future fertility rates for countries with total fertility above 2,1 children per woman are modeled based on the historical fertility declines observed in other countries. • Unlike previous assumptions, countries currently exceeding 2,1 children per woman are not restricted to this level; their fertility rates can drop as low as 1,85 children per woman, which is considered the lowest limit in this scenario. • Unlike the 2000 Revision, which had varying target fertility rates, all low-fertility countries are now expected to converge to 1,85 children per woman by the end of the projection period. • Fertility rates in the high and low variant projections deviate by plus or minus 0,5 children from the medium variant rate, a change from the 0,4 child difference used for low-fertility countries in the 2000 Revision.
2004	2005	2005I	Fertility: • For countries where total fertility was below 1,85 children per woman during 2000–2005, projections start by continuing recent fertility trends and then increase linearly by 0,07 children per woman every five years. Mortality: • Additional mortality models to better reflect the varied historical experiences in the rise of life expectancy. New models for very slow and very fast changes in mortality were developed, complementing the existing slow, medium, and fast models.
2006	2005	2005II	Fertility Projections: • For countries where total fertility was below 1,85 children per woman during 2000-2005, fertility trends are first projected to continue as they have been, then expected to rise linearly by 0,05 children per woman every five years, a reduction from the 0,07 increase used in the 2004 Revision. These countries are not projected to necessarily reach a fertility level of 1,85 children per woman by 2045–2050.
2008	2010	2010I	The starting year for the projections shifted from 2005 to 2010.
2010	2010	2010II	Fertility: • A new stochastic model for projecting fertility. • The medium-fertility variant represents the median of 100,000 projected trajectories for various countries. • Unlike previous versions that used a stabilization level of 1,85 children per woman, the 2010 Revision sets this level at 2,1 children per woman, acknowledging that many countries may only reach this level by the end of the century or beyond. Life Expectancy: • A new version of the model life tables had been developed, extending life expectancy projections from 75 to 100 years. This update also enhanced consistency with the mortality experiences of countries with the highest life expectancies, as per the Human Mortality Database.
2012	2010	2010III	Fertility Projections: • Continued using the stochastic model from 2010 with enhancements, incorporating a Bayesian hierarchical model for low-fertility countries to make projections more data-driven and specific to each country. • The median fertility estimate is based on 60 000 country-specific trajectories. Life Expectancy Projections: • Introduced two new models for countries not heavily impacted by HIV/AIDS. The first model uses a Bayesian hierarchical approach to assess improvements in female mortality based on life expectancy levels. The second model projects the gender gap in life expectancy. • The medium mortality estimate reflects the median of 100,000 sex-specific projected trajectories.
2015	2015	2015I	Fertility Projection: • A more structured method to project age-specific fertility patterns, enhancing precision and consistency. Life Expectancy Gender Gap:

			• Updated the model parameters for projecting the gender gap in life expectancy, raising the threshold value from 83 years in 2012 to 86 years in the 2015 Revision. Net Migration Assumptions: • Adjusted its approach assuming a gradual decline to 50% of its 2045-2050 level by 2095-2100, compared to an assumption of zero migration by 2095-2100 in the 2012 Revision.
2017	2015	2015II	Momentum Variant: • To demonstrate the influence of current age structures on future population dynamics, combining attributes of three existing scenarios: immediate replacement-level fertility, unchanged mortality, and zero net migration. Modified Probabilistic Models for Life Expectancy: • Adjustments were made to the parameters of probabilistic models projecting life expectancy at birth, enhancing the fit and consistency of projections and reducing the need for country-specific adjustments compared to previous revisions. Enhanced Sex Gap Projections: • Historical data pre-1950 were utilized to refine projections of the gender gap in life expectancy, particularly benefiting countries in earlier stages of mortality transition.
2019	2020	2020	International Migration: • It was projected that net international migration would maintain its 2045-2050 levels through the end of the 21st century, contrasting with the previous revision which anticipated a reduction to half those levels by 2095-2100.

Source: UN WPP, 1994-2019 revisions

Chapter 5

Error Measures

A number of standards have been used: the mean error (ME), the mean absolute error (MAE), the root mean square error (RMSE), the mean percentage error (MPE), the mean absolute percentage error (MAPE), and the root mean square percentage error (RMSPE) (Koutsandreas et al. 2022).

Forecast error, E, is defined as the difference between the projected value and the observed value:

$$E = (Projected - Observed)$$

Positive errors indicate optimistic forecasts, while negative error values suggest optimistic forecasts. The most popular scale-dependent measures are the ME, MAE, MSE, and RMSE defined as follows:

$$ME = \frac{\sum_{i=1}^{n}(Projected - Observed)}{n}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|Projected - Observed|$$

$$MSE = \frac{\sum_{i=1}^{n}(Projected - Observed)^2}{n}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(Projected - Observed)^2}{n}}$$

where:

n = the forecast horizon (Koutsandreas et al. 2022).

In all cases, smaller values imply better accuracy.

MAE is an appropriate measure for evaluating the ability of the forecasting method to specify the median of the future values of the series, while MSE for measuring the ability of the method to specify its mean (Kolassa 2016). RMSE has similar properties to MSE but has the advantage of being expressed on the same scale of the series, while also being less sensitive to outliers when its results are aggregated across multiple series (Hyndman and Koehler 2006). However, since both MSE and RMSE penalize more large errors than small ones, due to the squares used, MAE is less sensitive to outliers (Kolassa 2016). Although these measures are straightforward to compute, well defined, and interpretable, they are scale-dependent and, therefore, cannot be used for assessing forecasting accuracy across multiple time series. The RMSE gives much weight to big errors, whereas the ME and MAE give the same weight to all errors.

In order for the errors to become scale-independent to enable comparisons across different revisions and over time, a standard practice is to express the error measures as percentages:

$$\text{PE}=\left(\frac{Projected-Observed}{Observed}\right) \text{x } 100$$

A positive PE indicates an overestimation in the forecast, while a negative PE signifies an underestimation. By analyzing multiple forecasts, we can calculate the average errors. The MPE is used to gauge the average bias in these forecasts, with a positive MPE suggesting that the forecasts generally overestimate, and a negative MPE indicating a tendency to underestimate. For metrics that are not influenced by population size, such as the crude birth rate, the term 'mean error' is used, defined simply as the difference between the forecast and the observed value. Absolute errors, which do not consider the error's direction, measure the accuracy of forecasts by quantifying the deviation from actual values, regardless of whether the forecasts were too high or too low. The MPE, MAPE and the RMSPE are independent of the unit of measurement of the given variable. They can therefore be used to compare totally different situations.

Thus, the scale-dependent measures introduced earlier can be expressed in a percentage form, as follows:

$$\text{MPE} = \frac{1}{n}\sum_{i=1}^{n}\frac{(Projected-Observed)}{Observed} \times 100$$

$$\text{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\frac{|Projected-Observed|}{Observed} \times 100$$

$$\text{RMSPE} = \sqrt{\frac{\sum_{i=1}^{n}((Projected-Observed)/Observed)^2}{n}}$$

where:

n = the forecast horizon (Koutsandreas et al. 2022).

The main advantage of these measures, apart from enabling the evaluation of forecasting accuracy across multiple time series of different scales, is that they are easy to communicate, especially within businesses and organizations (Kolassa and Martin 2011), with MAPE being the most popular choice (Fildes and Goodwin 2007). However, they display several limitations that make their utilization problematic (Tofallis 2015). Since they depend on the division, they become infinite or undefined when observed value equals 0, which is typical in many applications such as inventory forecasting and intermittent demands. Moreover, the distribution of the above percentage measures becomes heavily skewed for observed values close to zero (Coleman & Swanson, 2004; Spiliotis et al., 2019; Swanson et al., 2000). Finally, measures that are based on percentage errors have the disadvantage that they put a different penalty on positive errors than on negative errors, thus being asymmetric (Hyndman and Koehler 2006). In order to mitigate the limitations discussed above, Makridakis (1993) introduced the symmetric APE, which divides E by the sum of projected and observed values instead of observed alone, as follows:

$$\text{sAPE} = \frac{|Projected - Observed|}{|Projected| + |Observed|} \times 200\%$$

$$\text{sMAPE} = \frac{1}{n}\sum_{i=1}^{n} \frac{|Projected - Observed|}{|Projected| + |Observed|}$$

where:

n = the forecast horizon

Note that the measure is defined even when observed equals zero, thus being more robust to values close to zero. However, despite its name, symmetric APE remains quite asymmetric, as it puts a heavier penalty on positive errors than on negative errors (Goodwin and Lawton 1999). In this regard, various modifications have been proposed for improving the properties of sAPE and other measures that build on percentage errors, until now, with mixed results (Hyndman and Koehler 2006). An interesting alternative to the measures that are based on percentage errors is the log of the accuracy ratio (logAR) (Tofallis 2015), expressed as follows:

$$\text{logAR} = \frac{1}{n}\sum_{i=1}^{n} \ln \left(\frac{Projected}{Observed}\right)^2$$

In contrast to MAPE, logAR is not biased, being also appropriate for specifying the geometric mean of the future values of the series. However, the measure becomes infinite or undefined when the observed value is equal to zero or when $\frac{Projected}{Observed}$ equals zero. Note that all the measures presented in this section display two major advantages. First, they allow pairing, meaning that each forecast is directly compared to the corresponding actual value. Hence, they are intuitive and can meaningfully handle trended data. Second, their calculation does not depend on a benchmark method, as is needed with relative errors and measures.

Theil's U statistics measures the difference between errors produced by a formal forecasting method and a naive alternative, such as the assumption that no change will occur. This statistic squares the errors so that large errors are given heavier weights than small errors.

Measuring the forecasting accuracy properly is a prerequisite for supporting decision making. However, till present, forecasters have not come into a general agreement about which measure should be preferred. For population projections accuracy measures are generally used solely to show how well (poorly) projections have performed as forecasts.

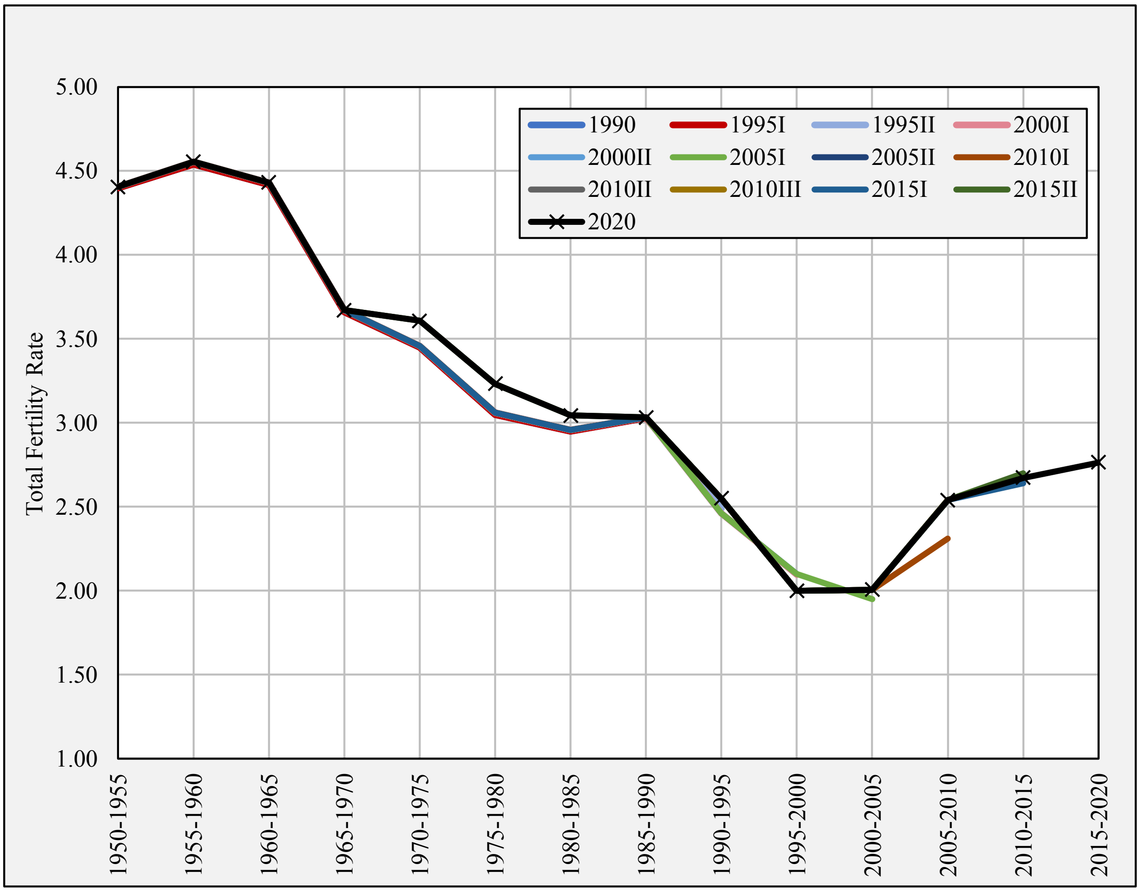
5.1 Fertility

The figure indicates that TFR values are virtually identical when compared across each revision. Between 1970 and 1980, the 2019 revision of the UN WPP revisited the estimated TFR for that decade, adjusting the figures upwards. Specifically, for the period 1970-1975, the TFR was increased by 0,17, and for 1975–1980, it was adjusted by an increase of 0,15.

Initially, the TFR was estimated at 4,41 for the period 1950–1955, peaking in the subsequent five-year span before gradually declining, with the lowest value recorded in the 1990s. The decline began between 1960-1965 and 1965-1970, attributed to the rapid pace of urbanization.

The urbanization of that period led to increased participation of women in social production, which subsequently contributed to a decline in birth rates (Portal "History of Kazakhstan" 2013).

Fig. 4: Total Fertility Rate Estimates, UN WPP, 1994–2019



Source: The author's calculations based on data from UN WPP 1994-2019 revisions.

Note: Revisions are abbreviated by the base year

Numerous field studies on urbanization and fertility trends reinforce the long-established conclusion from earlier demographic studies: fertility rates decline first and most rapidly in urban areas (Martine, Alves, and Cavenaghi 2013). As seen from the table, both the absolute number of the urban population and its share in the population structure of the republic increased. In 1970, compared to 1959, the urban population increased by almost 40%. The rapid growth of urbanization between 1959 and 1970 was due to significant migratory outflow from villages to cities, which can be explained by the substantial difference in the quality of life between urban and rural areas. The highest increase in migration was observed in the Guryeyv, Pavlodar, Kyzylorda, and Kostanay regions, which also had the highest intensity of migration from rural to urban areas (Portal "History of Kazakhstan 2013).

Tab. 3: Proportion of urban and rural population in Kazakhstan according to the 1959, 1970, and 1979 censuses

	1959		1970		1979		Index of change between 1959 and 1979
	thousands	%	thousands	%	thousands	%	(1959 = 100%)
Urban	4079,3	43,9	6745,4	51,8	7912,5	53,9	194,0
Rural	5203,9	56,1	6457,5	48,2	6775,8	46,1	130,2

Source: The author's calculations based on data from the Bureau of National Statistics of the Republic of Kazakhstan.

The state social security system, which was operating since the early 1960s as part of the virgin and fallow lands development program, supported the high birth rates in rural families by covering significant costs associated with childbearing. The rural lifestyle and the availability of subsidiary farms also encouraged larger families. However, the economic crisis following independence led to reduced support for rural areas and large families, narrowing the birth rate gap between urban and rural populations (Alekseenko 2019). The primary reason for the decline in the birth rates among the Kazakh population was the reduction in state financial support for large, typically rural, families.

An increase in TFR occurred between 1980–1985 and 1985–1990 was influenced by policies in the USSR. Specifically, the system of family assistance, established in the early 1980s through Resolution No. 235 and the "Main Directions of Economic and Social Development of the USSR for 1981-1985," had a significant impact on birth rates (Litvinets 2006). These policies represented critical steps in demographic strategy, reflecting a shift in attitudes towards fertility and family relations, and were discussed extensively among Soviet demographers.

The low TFR in the 1990s can be attributed to the economic crisis that followed after the dissolution of the USSR. From 2005 onwards, the TFR began to rise again.

Tab. 4: Errors in projected TFR by base year, projection period, and projection duration

	1990–1995	1995–2000	2000–2005	2005–2010	2010–2015	2015–2020		ME	MAE
1990	–0,05	0,37	0,22	–0,44	–0,57	–0,66		–0,19	0,39
1995I		0,30	0,09	–0,44	–0,57	–0,66		–0,26	0,41
1995II		0,30	0,09	–0,44	–0,57	–0,66		–0,26	0,41
2000I			–0,06	–0,64	–0,77	–0,86		–0,58	0,58
2000II			–0,06	–0,69	–0,82	–0,91		–0,62	0,62
2005I				–0,68	–0,82	–0,91		–0,80	0,80
2005II				–0,23	–0,47	–0,65		–0,45	0,45
2010I					–0,47	–0,65		–0,56	0,56
2010II					–0,19	–0,33		–0,26	0,26
2010III					–0,23	–0,40		–0,32	0,32
2015I						–0,23		–0,23	0,23
2015II						–0,19		–0,19	0,19
								–0,39	**0,43**

ME	–0,05	0,32	0,06	–0,51	–0,55	–0,59	**–0,22**		
MAE	0,05	0,32	0,10	0,51	0,55	0,59	**0,35**		
Duration in years	0–5	5–10	10–15	15–20	20–25	25–30			
ME	–0,15	–0,35	–0,54	–0,67	–0,63	–0,66	**–0,50**		
MAE	0,25	0,46	0,61	0,67	0,63	0,66	**0,55**		

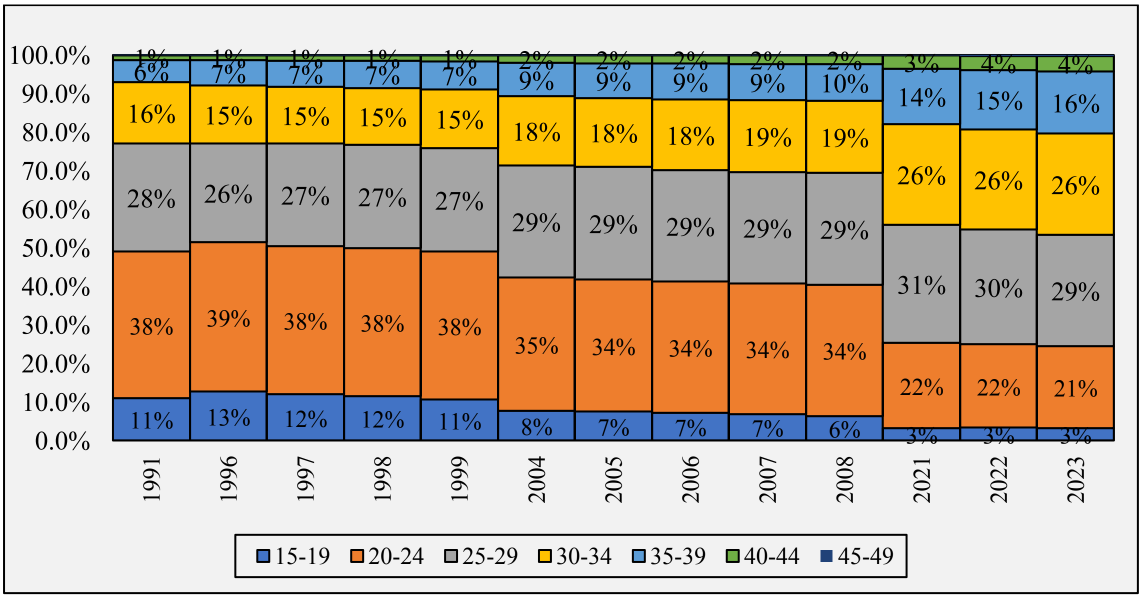
Source: The author's calculations based on data from UN WPP 1994-2019 revisions.
Note: The observed value represents the data from UN WPP 2019 revision.

The projections for TFR that used 1990 and 1995 as base years overestimated TFR levels for the 1990s. This overestimation can be attributed to several factors. Firstly, the collapse of the Soviet Union led to socio-economic instability, which impacted birth rates. Secondly, the fertility-promoting policies and anti-alcohol campaigns of the late 1980s resulted in many families having children earlier than planned, leading to a subsequent natural decline in birth rates (Bekkhozhayeva 2004). Starting from 2005, the TFR was projected to reach the replacement level of 2,1, which explains why these projections underestimated the TFR.

All later projections underestimated the TFR, as indicated by the negative errors observed. Notably, the projections that began in the years 2000 and 2005 (UN WPP 2004 revision), exhibited the highest mean errors. This suggests that these projections anticipated the lowest fertility scenarios, projecting TFRs below the replacement level.

The steady increase in the birth rate, which has been observed in Kazakhstan since the second half of the 2000s, was due to an increase in the number of women who entered childbearing age. These women, who were born during the birth rate surge of the mid-1980s, were at the most favorable age for childbearing — 20-29 years old. This age group accounted for about 63% of the children born during the year.

Fig. 5: Distribution of births by age of mothers



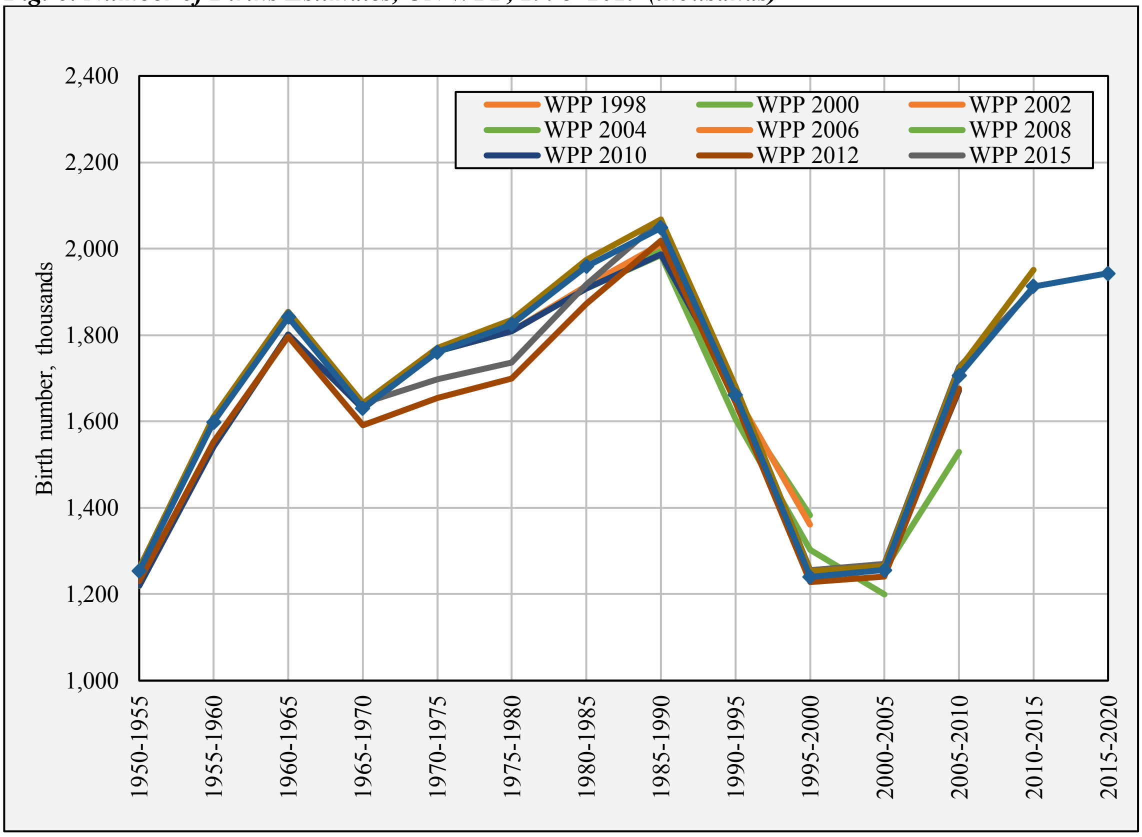
Source: The author's calculations based on data from UN WPP 1998-2019 revisions.

Since 2003, there has been a steady decrease in the teenage birth rate and a simultaneous increase in the number of births among women over the age of 30.

When examining the errors over different periods, it is evident that all projections consistently underestimated the TFR from 2005 onwards. Specifically, the average errors for the periods 2005–2010, 2010–2015, and 2015–2020 were –0,51, –0,55, and –0,59, respectively.

Projection accuracy decreases and bias increases with longer projection durations. Short-term projections (0–5 years) have lower bias (ME: –0,15) and moderate accuracy (MAE: 0,25), while extended projections (15–20 years) showed the highest bias (ME: –0,67) and lowest accuracy (MAE: 0,67).

Fig. 6: Number of Births Estimates, UN WPP, 1998–2019 (thousands)



Source: The author's calculations based on data from UN WPP 1998-2019 revisions.

The graph above shows that the estimated values of births did not change significantly across different revisions. Notably, the revisions from 2012 and 2015 estimated a lower number of births for the period between 1965 and 1980 compared to earlier revisions. However, from the 2017 revision onwards, the estimates returned to values similar to those of previous years.

Tab. 5: Errors in projected birth number by base year, projection period, and projection duration

	1995–2000	2000–2005	2005–2010	2010–2015	2015–2020		MPE	MAPE
1995II	21,04	10,68	-16,12	-25,63	-30,07		-8,02	20,71
2000I		4,26	-23,69	-33,52	-39,39		-23,08	25,21
2000II		0,16	-27,98	-36,62	-42,91		-26,84	26,92

2005I			-30,34	-37,48	-42,14		-36,65	36,65
2005II			-10,55	-22,11	-29,71		-20,79	20,79
2010I				-21,59	-28,81		-25,20	25,20
2010II				-11,17	-15,48		-13,32	13,32
2010III				-12,42	-17,42		-14,92	14,92
2015I					-6,97		-6,97	6,97
2015II					-5,47		-5,47	5,47
							-18,13	**19,62**
MPE	21,04	5,03	-21,74	-25,07	-25,84	**-9,31**		
MAPE	21,04	5,03	21,74	25,07	25,84	**19,74**		
Duration in years	0-5	5-10	10-15	15-20	20-25			
MPE	-7,31	-20,29	-31,62	-35,97	-30,07	**-25,05**		
MAPE	12,40	22,96	31,62	35,97	30,07	**26,60**		

Source: The author's calculations based on data from UN WPP 1998-2019 revisions.

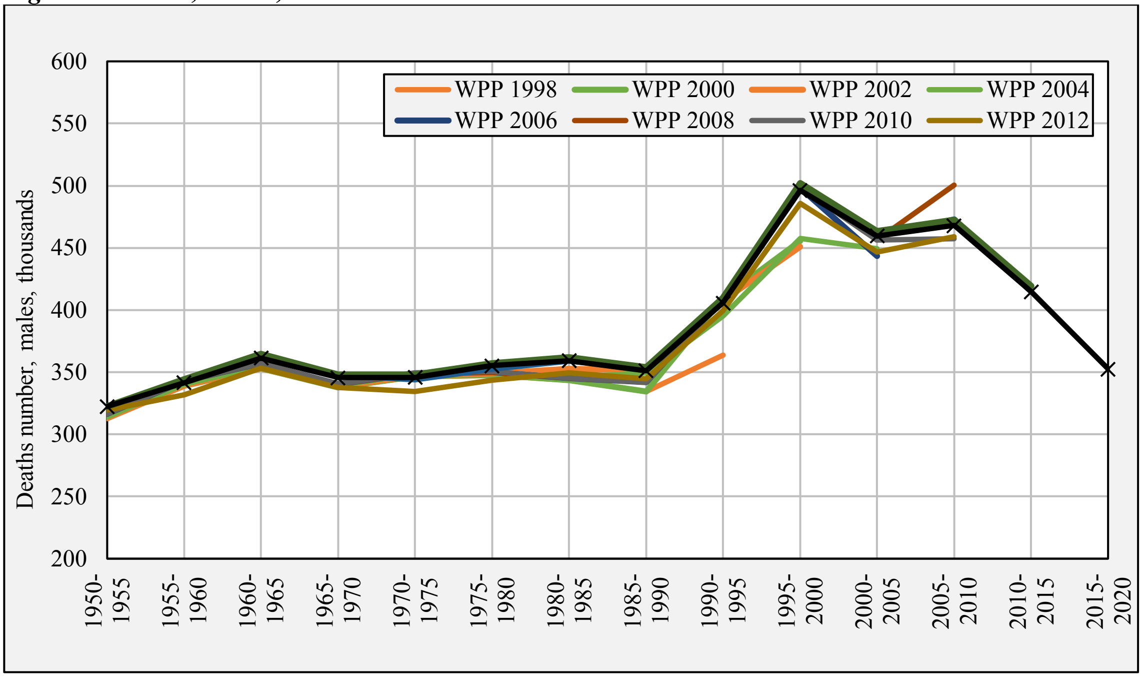
Note: The observed value represents the data from UN WPP 2019 revision.

The table reveals that between 1995 and 2005, the MPE across various revisions was positive, indicating an overestimation of birth numbers. Specifically, the UN 1998 projection starting in 1995 had an MPE of 21,04% at the base year alone. In contrast, all projection revisions beginning from the year 2005 showed a negative MPE at the base year, suggesting an underestimation of births. For instance, the projection beginning in 2005 (UN WPP 2004) underestimated the number of births at the base year by 30,34%. Overall, this particular projection recorded an MPE of -36.65% across all periods.

5.2 Mortality

Like the births estimates, the death estimates across revisions show no drastic differences. Generally, projections tended to underestimate the number of deaths. For example, the UN 1998 projection underestimated the death count by 41980 for the period 1990–1995. Similarly, the UN 2000, UN 2002, and UN 2004 projections each underestimated the number of deaths by around 40000 for the period 1995–2000.

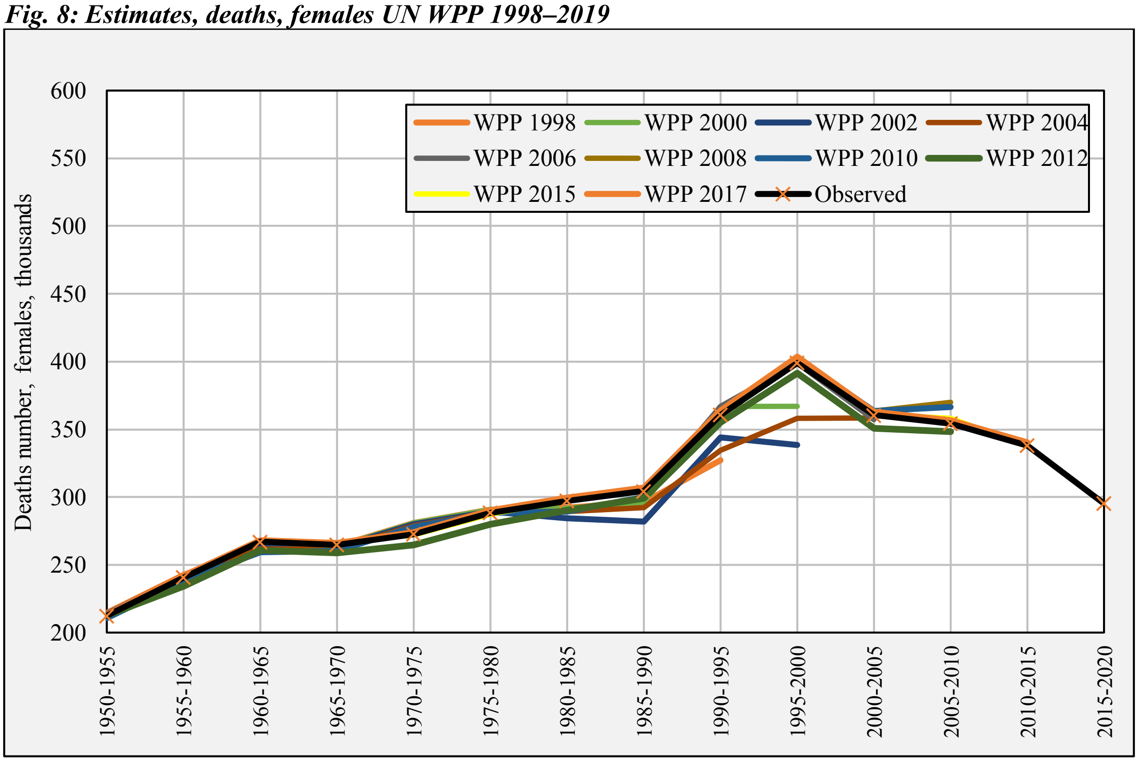
Fig. 7: Estimates, deaths, males UN WPP 1998-2019



Source: The author's calculations based on data from UN WPP 1998-2019 revisions.

Note: The observed value represents the data from UN WPP 2019 revision.

Fig. 8: Estimates, deaths, females UN WPP 1998–2019



Source: The author's calculations based on data from UN WPP 1998-2019 revisions.

Note: The observed value represents the data from UN WPP 2019 revision.

From 1950 to around 1980, all the estimation lines and the observed line remained relatively close together, indicating that estimations were in line with observed death numbers.

From 1990 onwards, the observed line shows a sharper increase in death numbers compared to the relatively flatter estimations, indicating a consistent underestimation of death numbers by the projections. This divergence is most pronounced in the periods 1990–1995 and 1995–2000, after which the estimations begin to more closely align with the observed values again, although they still underestimate the number of deaths to some extent.

From 2005–2010 and onwards, the estimations and observed values appear to align more closely, suggesting an improvement in the accuracy of the death number projections in the later years. The lines for UN 2015 and UN 2017 appear to be the closest to the observed data in the most recent periods.

The rise in mortality during the latter half of the 1990s and the early 2000s can be primarily attributed to the economy decline, that led to a reduction in healthcare expenditures and resulted in a healthcare crisis continually threatened by severe financial cutbacks (DHS 1995). By 1996, the healthcare budget had dropped to 1,1 percent of the Gross National Product, a stark contrast to the 6–10 percent typically spent on healthcare in most developed countries. Simultaneously, Kazakhstan's GNP had also decreased dramatically in the preceding years, further straining the budget (UNDP 1995). Since 1990 Kazakhstan's healthcare system had become increasingly decentralized and challenging to manage. Due to insufficient funding, certain regions reduced the number of hospital beds and the availability of essential drugs and medical equipment. Physicians in Kazakhstan were paid less on average than factory workers, and hospitals, along with other healthcare facilities, were in poor condition, often lacking sanitary conditions, running water, and electricity (Sharmanov et al. 1996).

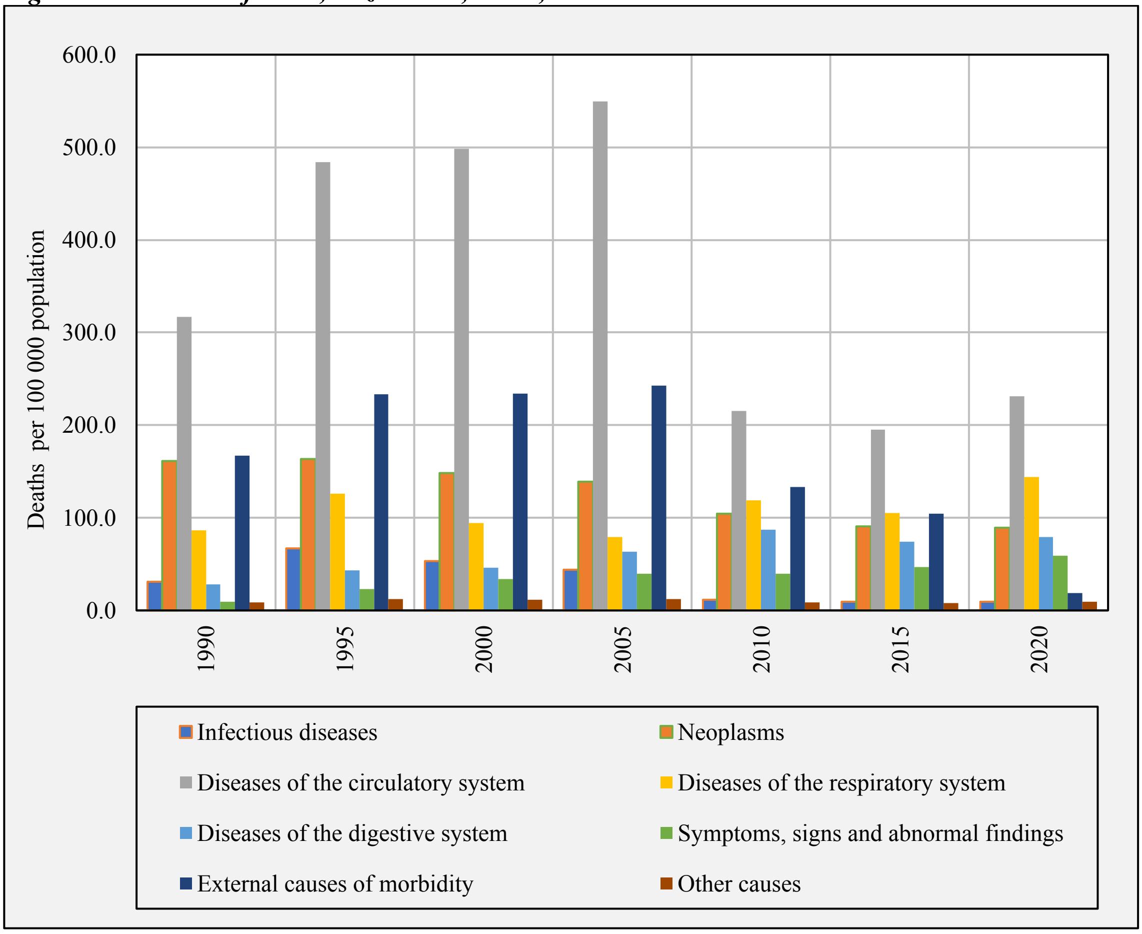
The crude death rate in Kazakhstan had risen from 7,7 deaths per 1,000 population in 1990 to 10,1 in 1995. Average life expectancy at birth had decreased from 68,6 years (63,8 for men and 73,1 for women) in 1990 to 66,8 (60,7 for men and 71,1 for women) in 1994 (DHS 1995). The primary causes of death in Kazakhstan were cardiovascular diseases, cancer, and respiratory diseases, with 484 134, and 93 deaths per 100 000 population, respectively in 1995. The overall morbidity rate from tuberculosis in 1995 was 271,1 per 100 000 population, making it the highest in Central Asia and one of the highest in the world. Drug-resistant forms of tuberculosis had become more common, leading to higher mortality and disability rates (DHS 1995).

The spike in mortality rates was particularly driven by an increase in circulatory system diseases (Fig.8 and Fig.9). This period, especially between 1995 and 2007, saw the mortality rate climb predominantly within the male population, contributing significantly to the overall increase (Bureau of National Statistics 2022). However, post-2008, the mortality gap between men and women has been gradually closing. In the early 1990s, women experienced higher mortality rates from circulatory system diseases compared to men, but this trend reversed by 2015, with men exhibiting higher mortality rates from these diseases in subsequent years, including 2019 and 2020. Additionally, mortality from external causes among men escalated from 167 per 100 000 in 1990 to 242 per 100 000 in 2005, with the mortality rate for men from these causes being three to four times higher than that for women (Bureau of National Statistics 2022).

Many of Kazakhstan's health problems also resulted from worsening environmental conditions. The health community had raised alarms about the significant increase in malignant neoplasms and genetic and mental disorders caused by radioactive contamination near the

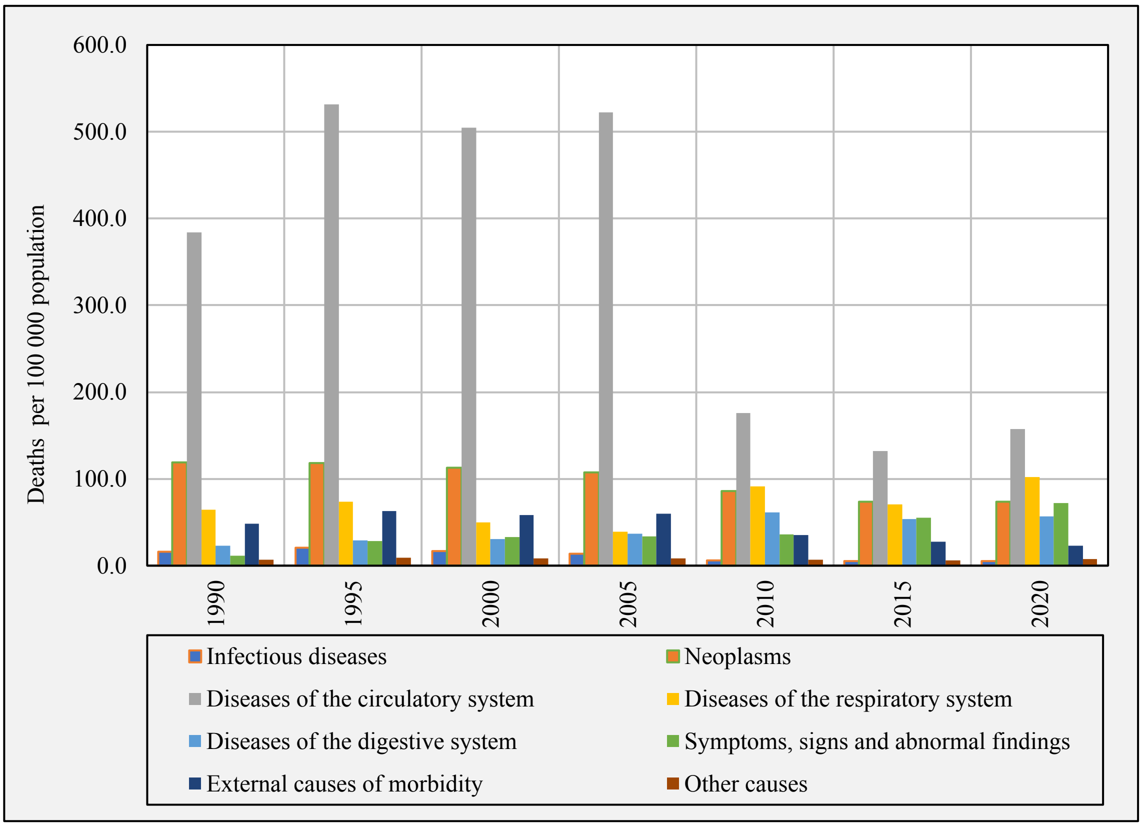
Semipalatinsk nuclear bomb testing zone and agro-chemical pollution in the ecologically devastated Aral Sea area. Besides environmental factors, behaviors such as heavy smoking, excessive alcohol consumption, and a high-fat diet had also played a substantial role in the declining health of Kazakhstan's population. Nutrition-related diseases, particularly those linked to malnutrition and micronutrient deficiencies, were a major public health concern, as they were key predisposing factors for infectious diseases and underlying causes of many noncommunicable diseases. Iron deficiency anemia had been recognized as a significant health problem in Kazakhstan for decades (DHS 1995).

Fig. 9: Main causes of death, Kazakhstan, males, 1990–2020



Source: The author's calculations based on data from the Bureau of National Statistics.
Notes: Other causes of death are presented in deaths per 1000 population.

Fig. 10: Main causes of death, Kazakhstan, females, 1990-2020



Source: The author's calculations based on data from the Bureau of National Statistics.
Notes: Other causes of death are presented in deaths per 1000 population.

The analysis of errors in the WPP estimates of the number of deaths for male population shows sizeable variability in accuracy across different periods. Early periods, such as 1950–1955 and 1980–1985, show substantial underestimations, with MPE of -15,64%, indicating notable inaccuracies. In contrast, more recent periods, especially from 2005 onwards, demonstrate improved accuracy, with some overestimations but generally lower error magnitudes. The MAE for the earlier periods ranges from 5,23% to 14,95%, while later periods show reduced values, reflecting increased reliability of estimates.

Similar to the analysis for the male population, the examination of errors in the estimates of the number of deaths among the female population reveals variability in accuracy across different periods and WPP revisions. Early periods, such as 1960–1965 and 1980–1985, demonstrate underestimations in the number of deaths, with the MPE of -13.06% and MAE of up to 6,53%. On the other hand, more recent periods, once again from 2005 onwards, show improved accuracy, with lower MPE and MAE values, which is indicative of more reliable estimates. For instance, the WPP 2017 revision demonstrates the highest accuracy with an MPE of 2,10% and an MAE of 2,10%.

Tab. 6 Errors in the estimations of the number of deaths, males (%)

Year	WPP 1998	WPP 2000	WPP 2002	WPP 2004	WPP 2006	WPP 2008	WPP 2010	WPP 2012	WPP 2015	WPP 2017		ME	MAE
1950–1955	-9,88	-8,84	-8,98	-8,56	-6,07	-6,07	-6,00	-2,96	-0,07	-0,07		-5,23	5,23
1955–1960	-2,80	-1,86	-1,84	-1,26	1,57	1,57	1,66	-10,03	2,75	2,75		-0,68	2,56
1960–1965	-8,97	-8,27	-8,00	-7,23	-4,73	-4,73	-4,65	-8,11	2,83	2,83		-4,46	5,49
1965–1970	-8,10	-7,73	-7,29	-6,24	-5,01	-5,01	-4,94	-7,61	2,49	2,49		-4,27	5,17
1970–1975	0,47	0,71	1,67	2,67	3,37	3,37	3,45	-11,14	-1,60	2,03		0,45	2,77
1975–1980	-7,79	-7,63	-6,18	-5,38	-5,03	-5,03	-4,93	-11,94	-2,51	1,91		-4,96	5,30
1980–1985	-15,64	-15,59	-6,52	-13,68	-14,30	-14,30	-14,11	-9,69	0,10	2,48		-9,20	9,67
1985–1990	-16,44	-16,44	0,39	-1,68	-9,29	-9,29	-8,90	-6,61	2,03	2,70		-5,78	6,71
1990–1995	-41,98	3,83	1,61	-10,82	3,32	2,79	2,94	-6,36	3,14	4,01		-3,41	7,35
1995–2000	x	-40,78	-45,36	-38,71	1,88	2,20	2,32	-10,51	2,12	5,63		-12,12	14,95
2000–2005	x	x	x	-10,23	-16,28	-3,39	-3,34	-12,99	-0,22	3,64		-5,35	6,26
2005–2010	x	x	x	x	x	32,56	-10,34	-8,82	5,19	4,42		3,83	10,22
2010–2015	x	x	x	x	x	x	x	x	0,15	4,48		1,54	1,54
ME	-12,38	-10,26	-8,05	-9,19	-4,60	-0,44	-3,90	-8,90	1,26	3,02			
MAE	12,45	11,17	8,78	9,68	6,44	7,53	5,63	8,90	1,94	3,04			

Source: The author's calculations based on data from UN WPP 1998-2019 revisions.

Note: The observed value represents the data from UN WPP 2019 revision.

Tab. 7: Errors in the estimations of the number of deaths, females (%).

Year	WPP 1998	WPP 2000	WPP 2002	WPP 2004	WPP 2006	WPP 2008	WPP 2010	WPP 2012	WPP 2015	WPP 2017		ME	MAE
1950-1955	2,08	1,04	1,18	0,76	-1,73	-1,73	-1,80	0,15	0,75	0,75		0,14	1,20
1955-1960	2,08	1,14	1,12	0,54	-2,29	-2,29	-2,38	-6,20	1,22	1,22		-0,59	2,05
1960-1965	-3,58	-4,28	-4,55	-5,32	-7,82	-7,82	-7,90	-6,14	1,03	1,03		-4,53	4,95
1965-1970	-1,22	-1,58	-2,02	-3,08	-4,31	-4,31	-4,37	-6,03	1,12	1,12		-2,47	2,92
1970-1975	8,31	8,07	7,12	6,11	5,41	5,41	5,33	-7,97	-0,79	1,51		3,85	5,60
1975-1980	2,65	2,49	1,03	0,23	-0,12	-0,12	-0,21	-8,41	-0,89	1,84		-0,15	1,80
1980-1985	-5,33	-5,39	-13,06	-8,26	-7,65	-7,65	-7,83	-7,11	0,88	2,15		-5,92	6,53
1985-1990	-8,16	-8,15	-22,38	-12,37	-4,17	-4,17	-4,56	-5,38	2,12	2,59		-6,46	7,40
1990-1995	-33,49	6,21	-16,68	-26,11	6,32	0,43	0,32	-5,45	2,84	3,37		-6,22	10,12
1995-2000	x	-32,17	-60,74	-41,15	-0,27	3,31	3,21	-7,57	2,41	4,59		-14,26	17,27
2000-2005	x	x	x	-2,11	-3,43	2,85	2,78	-9,94	0,39	2,75		-0,96	3,46
2005-2010	x	x	x	x	x	15,69	12,13	-6,04	3,72	2,13		5,53	7,94
2010-2015	x	x	x	x	x	x	x	x	-0,68	2,20		0,76	1,44
ME	-4,07	-3,26	-10,90	-8,25	-1,82	-0,03	-0,44	-6,34	1,09	2,10			
MAE	7,43	7,05	12,99	9,64	3,96	4,65	4,40	6,37	1,45	2,10			

Source: The author's calculations based on data from UN WPP 1998-2019 revisions.

Note: The observed value represents the data from UN WPP 2019 revision.

For the period 1995–2000, the projections significantly underestimated the number of female deaths. Specifically, the earlier revisions such as WPP 2000, WPP 2002, and WPP 2004, fell short of the actual female death count by 32170, 60740, and 41150 deaths respectively.

Tab. 8: Errors in projected deaths by base year, projection period, and projection duration, males, %

	1995-2000	2000-2005	2005-2010	2010-2015	2015-2020		MPE	MAPE
1995II	-24,59	-22,80	-23,59	-12,06	8,26		-14,96	18,26
2000I		-2,90	-8,95	-0,23	14,80		0,68	6,72
2000II		-11,13	-16,03	-8,60	5,11		-7,66	10,22
2005I			-4,01	8,69	25,98		10,22	12,89
2005II			-8,49	-2,20	11,92		0,41	7,53
2010I				17,04	31,26		24,15	24,15
2010II				4,32	22,76		13,54	13,54
2010III				14,16	37,86		26,01	26,01
2015I					27,72		27,72	27,72
2015II					26,78		26,78	26,78
							10,69	**17,38**
MPE	-24,59	-12,28	-12,21	2,64	21,24	**-5,04**		
MAPE	24,59	12,28	12,21	8,41	21,24	**15,75**		
Duration in years	0-5	5-10	10-15	15-20	20-25			
MPE	3,89	6,32	1,10	2,62	8,26	**4,44**		
MAPE	14,11	18,82	14,06	10,66	8,26	**13,18**		

Source: The author's calculations based on data from UN WPP 1998-2019 revisions.
Note: The observed value represents the data from UN WPP 2019 revision.

Tab. 9: Errors in projected deaths by base year, projection period, and projection duration, females, %

	1995-2000	2000-2005	2005-2010	2010-2015	2015-2020		MPE	MAPE
1995II	-18,95	-12,02	-9,09	-2,11	16,45		-5,15	11,73
2000I		-0,27	-0,64	3,77	20,89		5,94	6,39
2000II		-10,04	-7,43	-1,02	16,83		-0,42	8,83
2005I			2,57	9,27	26,69		12,84	12,84
2005II			-1,03	1,79	17,41		6,06	6,74
2010I				11,91	29,85		20,88	20,88
2010II				5,49	23,67		14,58	14,58
2010III				4,24	24,98		14,61	14,61
2015I					23,93		23,93	23,93
2015II					22,53		22,53	22,53
							11,58	**14,31**
MPE	-18,95	-7,45	-3,13	4,17	22,32	**-0,61**		
MAPE	18,95	7,45	4,15	4,95	22,32	**11,56**		

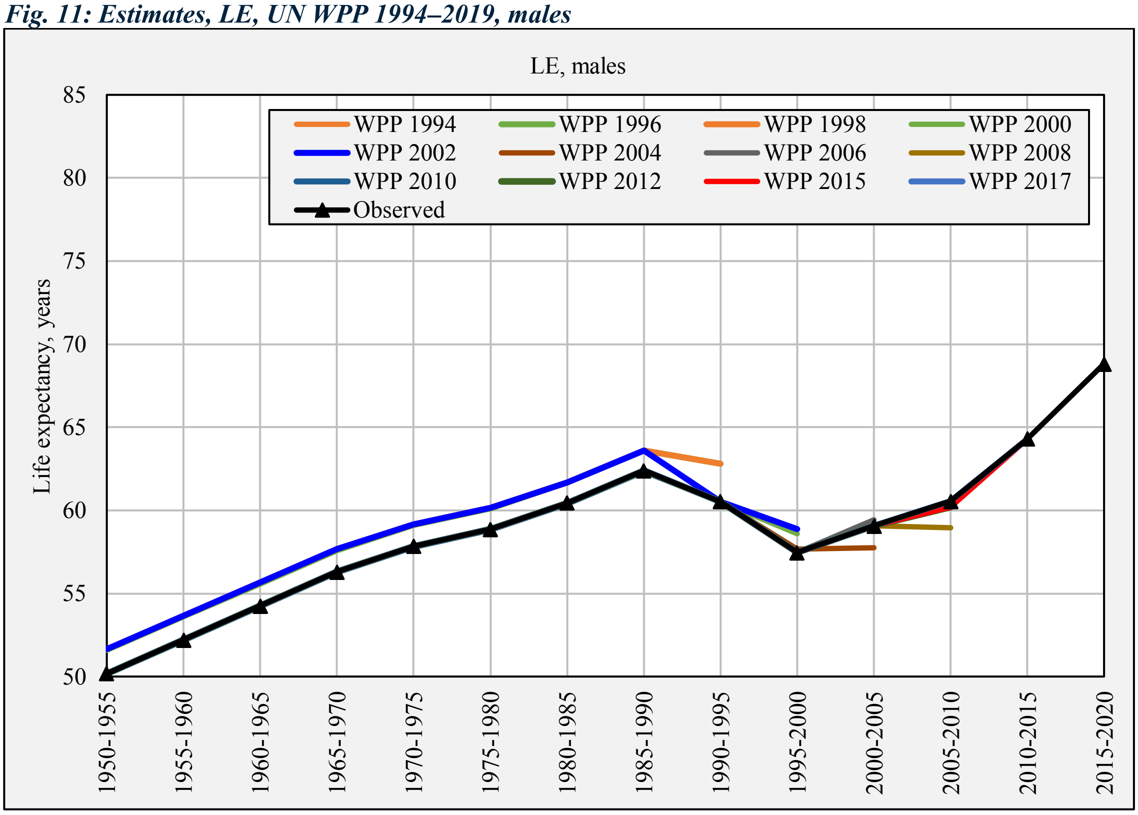
Duration in years	0-5	5-10	10-15	15-20	20-25			
MPE	4,04	8,68	7,55	11,87	16,45	**9,72**		
MAPE	10,10	13,71	11,60	13,28	16,45	**13,03**		

Source: The author's calculations based on data from UN WPP 1998-2019 revisions.

Note: The observed value represents the data from UN WPP 2019 revision.

From the tables it can be seen that the earlier projections with jump-off year 1995 and 2000 underestimated the number of deaths for both female and male population. However for the later projections with the jump-off year 2010 and 2015, the projections were overestimating the number of deaths for both males and females. In general MPE by the duration of projections were all positive, indicating that the forecasters were generally picking the high mortality scenarios (probably influenced by the high number of deaths during the first decade of independency).

Fig. 11: Estimates, LE, UN WPP 1994–2019, males



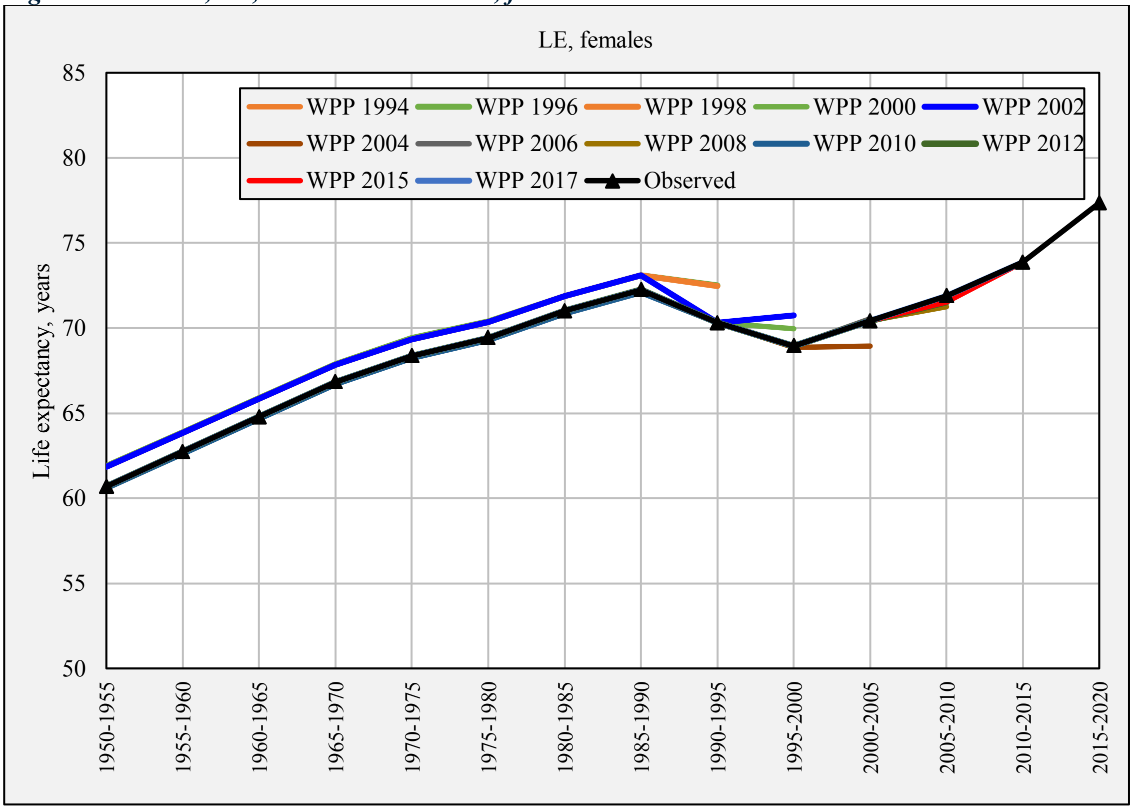
Source: The author's calculations based on data from UN WPP 1994-2019 revisions.

Both graphs display a general upward trend in life expectancy over time, reflecting the global improvement in health and longevity. The female life expectancy graph shows a steady increase, with the observed values closely tracking or slightly exceeding the estimates. This suggests that the WPP revisions for females were relatively accurate or slightly conservative.

In contrast, the male life expectancy graph exhibits more fluctuation, particularly in the periods around 1990-2005, where there's a noticeable dip before the trend continues upward. The observed life expectancy for males tends to be slightly lower than the estimates in earlier revisions, but by the WPP 2017 revision, the estimates and observed data converge more closely.

Comparing the two, the female life expectancy is consistently higher than that of males, which is a common trend in demographic studies, where females typically have a longer life expectancy than males. The disparities between the estimates and observed data across revisions appear to narrow over time for both genders, indicating that the accuracy of life expectancy projection has improved in more recent WPP revisions.

Fig. 12: Estimates, LE, UN WPP 1994–2019, females



Source: The author's calculations based on data from UN WPP 1994-2019 revisions.

Tab. 10: Errors in Projected values of LE, males (%)

	1990-1995	1995-2000	2000-2005	2005-2010	2010-2015	2015-2020		MPE	MAPE
1990	4,47	9,04	8,93	8,45	5,71	1,98		6,43	6,43
1995I		5,34	5,73	5,75	3,51	-0,02		4,06	4,07
1995II		5,34	5,73	5,75	3,51	-0,02		4,06	4,07
2000I			0,53	1,05	-0,69	-3,22		-0,58	1,37
2000II			1,79	2,31	0,57	-1,96		0,68	1,66
2005I				-1,82	-4,42	-7,55		-4,60	4,60
2005II				1,09	-0,4	-3,09		-0,80	1,53
2010I					-4,2	-7,01		-5,61	5,61
2010II					-2	-4,98		-3,49	3,49
2010III					-3,35	-7,16		-5,26	5,26
2015I						-3,95		-3,95	3,95
2015II						-3,63		-3,63	3,63
								-1,06	**3,80**
MPE	4,47	6,57	4,54	3,23	-0,18	-3,38	**2,54**		
MAPE	4,47	6,57	4,54	3,75	2,84	3,71	**4,31**		

Duration in years	0-5	5-10	10-15	15-20	20-25	25-30			
MPE	-0,03	-0,01	1,38	2,06	1,89	1,98	**1,21**		
MAPE	3,13	4,78	4,62	4,13	1,92	1,98	**3,43**		

Source: The author's calculations based on data from UN WPP 1994-2019 revisions.

Note: The observed value represents the data from UN WPP 2019 revision.

Across the different periods, the MPEs start off as positive in the early periods (1990-1995 to 2005-2010), indicating a trend of overestimating male life expectancy. As we move towards the later periods (2010-2015 and 2015-2020), MPEs become negative, showing a trend of underestimation. The MAPEs are consistently positive, as they are absolute values, and show the magnitude of errors irrespective of direction. The errors peak during the 1995-2000 period and decrease towards the most recent periods, albeit with some fluctuations.

Looking at the MPEs and MAPEs down the columns, we can assess the accuracy of projections made in different base years. The 1990 base year, being influenced by optimistic views of longevity, shows a high degree of overestimation (MPE of 6.43%), which decreases slightly in the subsequent 1995I and 1995II base years. From the 2000I base year onward, the projections display a mixed pattern, shifting from slight overestimation to underestimation in more recent periods. Notably, projections from 2005I and onward show a pronounced underestimation of life expectancy.

When analyzing by projection duration, the earlier durations (0-5 and 5-10 years) show nearly zero MPE, indicating high accuracy. Errors become more pronounced with longer projection durations (10-15 years onwards), which is typical as longer-term projections are subject to greater uncertainty.

Tab. 11: Errors in Projected values of LE, females (%)

	1990-1995	1995-2000	2000-2005	2005-2010	2010-2015	2015-2020		MPE	MAPE
1990	3,6	6,05	5,38	4,7	3,53	0,55		3,97	3,97
1995I		3,55	2,98	2,6	1,53	-1,15		1,90	2,36
1995II		3,5	3,03	2,55	1,58	-1,1		1,91	2,35
2000I			0,29	0,01	-0,76	-3,24		-0,93	1,08
2000II			1,52	1,24	0,27	-2,21		0,20	1,31
2005I				-2,11	-3,11	-5,5		-3,57	3,57
2005II				0,5	0,12	-2,1		-0,49	0,91
2010I					-1,78	-4,16		-2,97	2,97
2010II					-1,03	-3,49		-2,26	2,26
2010III					-1,61	-4,39		-3,00	3,00
2015I						-2,82		-2,82	2,82
2015II						-2,5		-2,50	2,50
								-0,88	**2,42**
MPE	3,60	4,37	2,64	1,36	-0,13	-2,68	**1,53**		
MAPE	3,60	4,37	2,64	1,96	1,53	2,77	**2,81**		

Duration in years	0-5	5-10	10-15	15-20	20-25	25-30			
MPE	0,09	-0,17	0,35	0,47	0,43	0,55	**0,29**		
MAPE	2,07	2,86	2,74	2,65	1,93	0,55	**2,13**		

Source: The author's calculations based on data from UN WPP 1994-2019 revisions.

Note: The observed value represents the data from UN WPP 2019 revision.

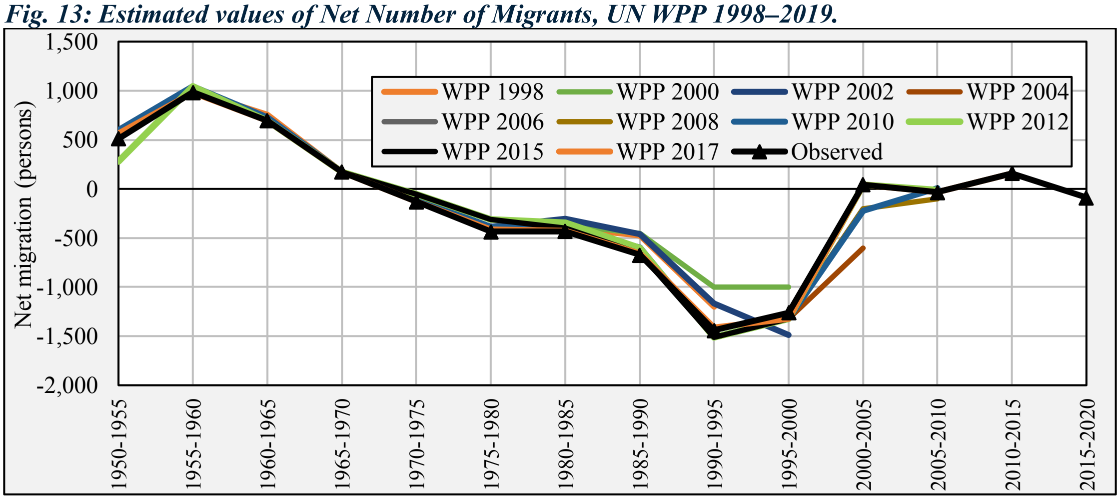
The projections for female life expectancy show an initial trend of overestimation then shifting to underestimation in more recent years. From 1990-1995 to 2005-2010, projections show a consistent overestimation of female life expectancy, with the MPE decreasing over these periods. The MPE starts at 3,6% in the 1990-1995 period and declines to 1,36% in the 2005-2010 period. In the 2010-2015 and 2015-2020 periods, this trend reverses to an underestimation, with MPEs of -0,13% and -2,68% respectively, indicating forecasters did not anticipate the actual improvements in female life expectancy.

The 1990 base year projection overestimates female life expectancy across all subsequent periods, with the overestimation gradually lessening by the 2015-2020 period. Projections starting in 1995 (1995I and 1995II) also overestimate life expectancy, though they show a transition to slight underestimation by the 2015-2020 period. Starting from the 2005 base year projections, there is a consistent pattern of underestimation, growing more pronounced by the 2015–2020 period.

Over shorter projection durations (0–5 years and 5–10 years), the MPEs are close to zero, suggesting that predictions for the near future were quite accurate. As the duration increases, the MPE becomes more variable, but the trend is not strictly linear, indicating that the accuracy of the projections varies and does not solely depend on the projection duration. The MAPE generally shows a moderate degree of error that does not exceed 3% suggesting that while the direction of error (over- or underestimation) varies, the magnitude of error remains within a similar range over time.

5.3 Migration

Fig. 13: Estimated values of Net Number of Migrants, UN WPP 1998–2019.



Source: The author's calculations based on data from UN WPP 1994-2019 revisions.

As it was mentioned previously, between 1955 and 1960, Kazakhstan experienced a substantial wave of immigration due to the virgin lands campaign and industrial development (Rybalkin 2017). However, by 1970, this trend began to reverse, with migration outflow starting to outweigh the influx. This outflow became more pronounced following the dissolution of the USSR. The migration trends during this period were driven by a complex interplay of socio-economic and political factors that directly impacted living standards. The dissolution of the Soviet Union aggravated economic troubles, characterized by a drop in production, rampant inflation, and rising unemployment, both apparent and concealed. These factors contributed to a declining standard of living and a crisis in social services, which, over time, led to a reduction in the number of people leaving the country and a subsequent decrease in the negative migration balance (Alimbekova and Zeynalova n.d).

Tab. 12: Errors in Projected values of Net Migration, UN WPP 1998–2019

	1995-2000	2000-2005	2005-2010	2010-2015	2015-2020		ME	MAE
1995II	-1088,00	-800,00	-400,00	-300,00	-200,00		-557,60	557,60
2000I		-800,00	-600,00	-350,00	-300,00		-512,50	512,50
2000II		-800,00	-743,00	-300,00	-300,00		-535,75	535,75
2005I			-400,00	-300,00	-300,00		-333,33	333,33
2005II			-200,00	-200,00	-200,00		-200,00	200,00
2010I				-100,00	-100,00		-100,00	100,00
2010II				-50,01	-50,01		-50,01	50,01
2010III				0,00	0,00		0,00	0,00
2015I					0,00		0,00	0,00
2015II					0,00		0,00	0,00
							-228,92	**228,92**
ME	-1088,00	-800,00	-468,60	-200,00	-145,00	**-540,32**		
MAE	1088,00	800,00	468,60	200,00	145,00	**540,32**		
Duration in years	0-5	5-10	10-15	15-20	20-25			
ME	-343,80	-349,13	-310,00	-300,00	-200,00	**-300,59**		
MAE	343,80	349,13	310,00	300,00	200,00	**300,59**		

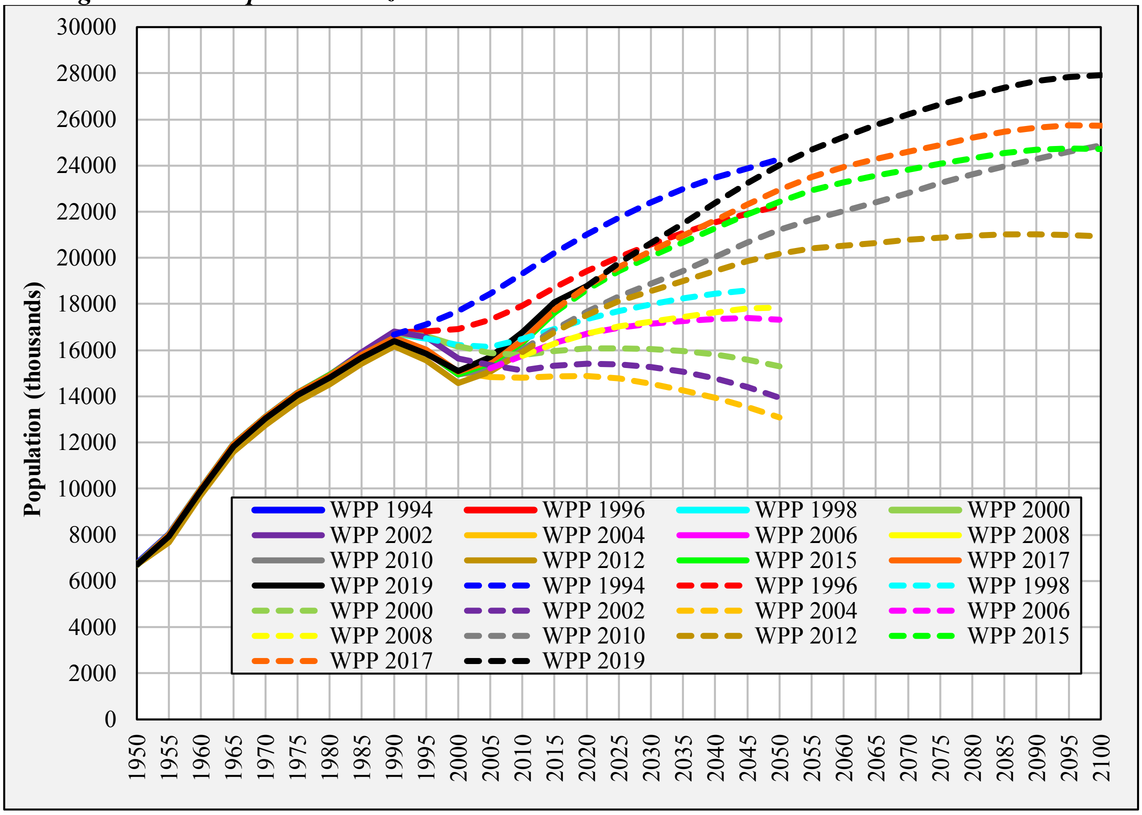
Source: The author's calculations based on data from UN WPP 1998–2019 revisions.

Note: The observed value represents the data from UN WPP 2019 revision.

The ME and MAE values indicate that projections have been consistently underestimating the net migration in Kazakhstan, but the extent of underestimation has decreased with each subsequent WPP revision. The projections for the 1995–2000 period started with an error of -1088, indicating a substantial underestimation of net outflow. The error decreases over time, with the 2005—2010 period showing an ME of -468,6, and further reducing to -200 by the 2010–2015 period. The 1995II base year projection had the largest absolute error of 557,6, indicating a large gap between projected and observed values. Subsequent base year forecasts show a trend of decreasing absolute error, with the 2000 base years having an MAE around 512–536, which decreases to 333 for the 2005I projection and even further to 200 for the 2005II projection.

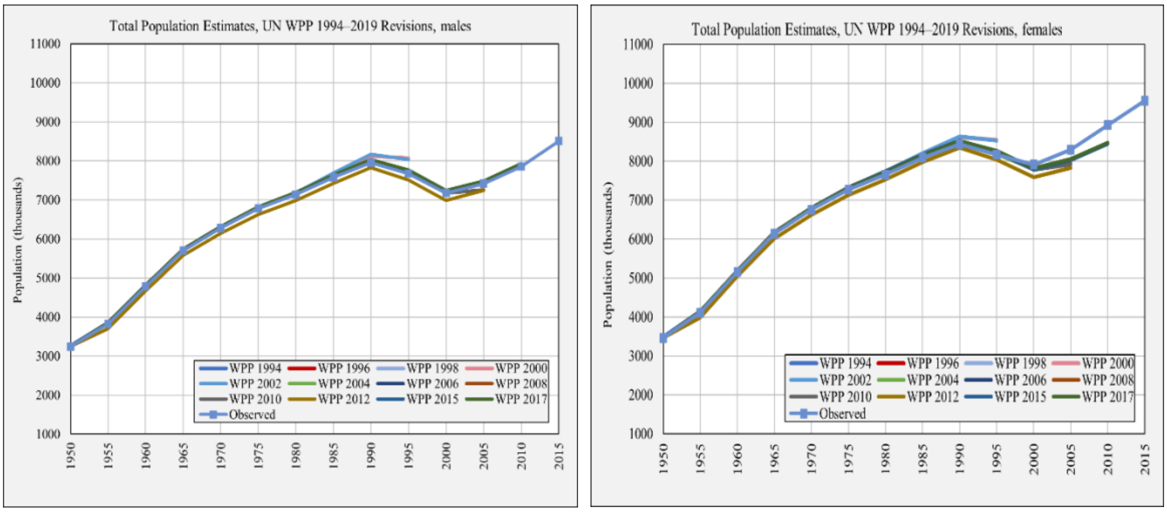
5.4 Total Population

Fig. 14: Total Population. Kazakhstan. Historical estimates and UN WPP 1994–2019



Source: The author's calculations based on data from UN WPP 1994–2019 revisions.

Fig. 15: Estimates of Total Population number by sex, UN WPP 1994–2019



Source: The author's calculations based on data from UN WPP 1994–2019 revisions.

Both graphs depict a consistent rise in population from 1950 to around 1990, followed by a decline throughout the 1990s. Post-2000, the population figures begin to rise again. In each graph, the estimates and observed data points align closely, demonstrating accurate estimations. Notable differences in the estimates occur between 1990 and 2005, where variations among the WPP revisions are more apparent.

Tab. 13: Errors in projected total population by base year, projection period, and projection duration, males, UN WPP 1994-2019, %

	1990	1995	2000	2005	2010	2015		MPE	MAPE
1990	1,70	8,52	20,27	21,56	20,59	16,64		14,88	14,88
1995I		6,54	14,74	13,66	11,28	7,46		10,74	10,74
1995II		4,61	10,01	6,02	2,47	-2,80		4,06	5,18
2000I			9,29	3,56	-2,78	-9,40		0,16	6,26
2000II			4,93	-0,73	-7,98	-13,95		-4,43	6,90
2005I				-4,28	-10,10	-16,85		-10,41	10,41
2005II				-1,96	-4,15	-8,45		-4,85	4,85
2010I					-4,57	-9,12		-6,84	6,84
2010II					-2,05	-4,64		-3,35	3,35
2010III					-2,31	-5,21		-3,76	3,76
2015I						0,05		0,05	0,05
2015II						1,02		1,02	1,02
								-0,23	**6,19**
MPE	1,70	6,56	11,85	5,40	0,04	-3,77	**3,63**		
MAPE	1,70	6,56	11,85	7,40	6,83	7,97	**7,05**		
Duration in years	0	5	10	15	20	25	30		
MPE	1,08	0,29	0,56	2,39	8,42	16,64	**4,90**		
MAPE	3,61	7,08	10,86	11,73	10,28	16,64	**10,03**		

Source: The author's calculations based on data from UN WPP 1994–2019 revisions.

Note: The observed value represents the data from UN WPP 2019 revision.

The table displays errors in projected total population values for males, segmented by base year and projection duration. Analyzing trends from the base year 1990 through to 2015, the MPE shifts from positive to negative, indicating a trend from initial overestimations to subsequent underestimations of population sizes. The earliest projections, beginning in 1990 and 1995I, show overestimations with MPEs of 14,88% and 10,74%, respectively. The most substantial underestimation occurs in the 2005I projection, with an MPE of -10,41%.

Further scrutiny of MPE and MAPE across different projection durations reveals that projection accuracy generally diminishes as the duration extends. Shorter-term projections (within 5–10 years from the base year) display lower absolute errors compared to long-term projections (20–25 years out), which exhibit the most significant peaks in error. Notably, the period from 1995 to 2005 recorded the highest errors, highlighting a period of particular challenge in accurately forecasting population changes. This pattern underscores the complexities and increasing difficulties of demographic forecasting as the prediction interval lengthens.

Tab. 14: Errors in projected total population by base year, projection period, and projection duration, females, UN WPP 1994-2019, %

	1990	1995	2000	2005	2010	2015		MPE	MAPE
1990	1,79	7,56	14,57	13,58	10,37	7,81		9,28	9,28
1995I		5,81	9,91	7,05	2,76	0,23		5,15	5,15
1995II		3,84	5,28	-0,22	-5,44	-9,45		-1,20	4,85
2000I			5,30	-1,21	-8,56	-13,63		-4,53	7,18
2000II			2,52	-3,54	-11,50	-16,17		-7,17	8,43
2005I				-6,87	-13,31	-18,31		-12,83	12,83
2005II				-4,29	-7,83	-10,91		-7,68	7,68
2010I					-7,52	-10,41		-8,96	8,96
2010II					-6,68	-8,15		-7,42	7,42
2010III					-7,63	-8,86		-8,25	8,25
2015I						-4,59		-4,59	4,59
2015II						-4,16		-4,16	4,16
								-4,36	**7,40**
MPE	1,79	5,74	7,51	0,64	-5,53	-8,05	**0,35**		
MAPE	1,79	5,74	7,51	5,25	8,16	9,39	**6,31**		
Duration in years	0	5	10	15	20	25	30		
MPE	-1,87	-3,06	-3,98	-3,78	0,39	7,81	**-0,75**		
MAPE	5,08	7,61	10,16	10,32	6,68	7,81	**7,94**		

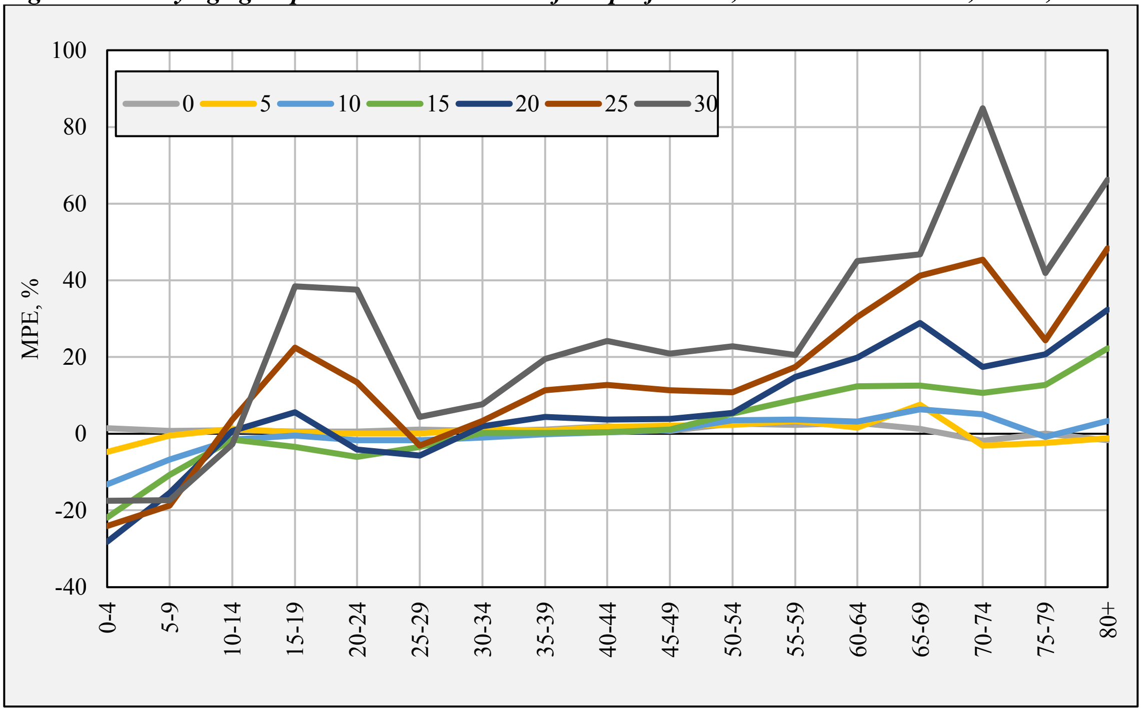
Source: The author’s calculations based on data from UN WPP 1994–2019 revisions.

Note: The observed value represents the data from UN WPP 2019 revision.

The table displays errors in projected total population figures for females, sorted by base year, projection period, and projection duration, highlighting variations in accuracy across different time frames. Initially, projections like those of 1990 and 1995I moderately overestimated population sizes, but starting from the 1995II projection, a consistent trend of underestimation began, particularly pronounced in projections with base years 2005 and 2010. This pattern indicates that early projections tended to overestimate population sizes up until 2000, after which underestimations became more common, especially noticeable from 2010 onwards. The MAPEs tend to rise with the length of the projection period, suggesting increasing uncertainty over longer durations. Notably, even at the base year, there is an error rate of 5,08%, indicating inaccuracies in the base population estimates.

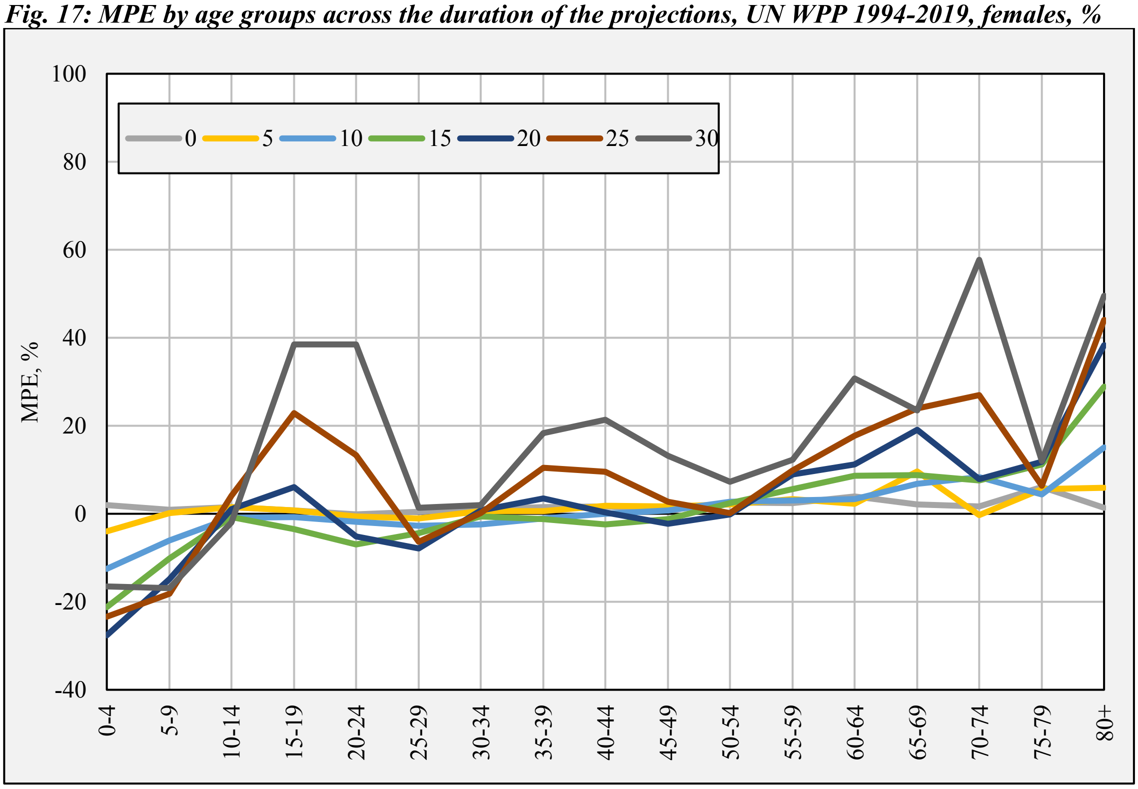
The graphs (Fig. 16 and Fig. 17) reveal that for short projection durations (0, 5, 10 years), the MPEs are very small, nearly approaching zero. However, as the projection duration increases, the magnitude of the MPEs also increases. For both males and females, there is a consistent overestimation in the older age groups and in the age groups of 15–20 and 20–24. Conversely, there is a consistent underestimation in the youngest age groups.

Fig. 16: MPE by age groups across the duration of the projections, UN WPP 1994-2019, males, %



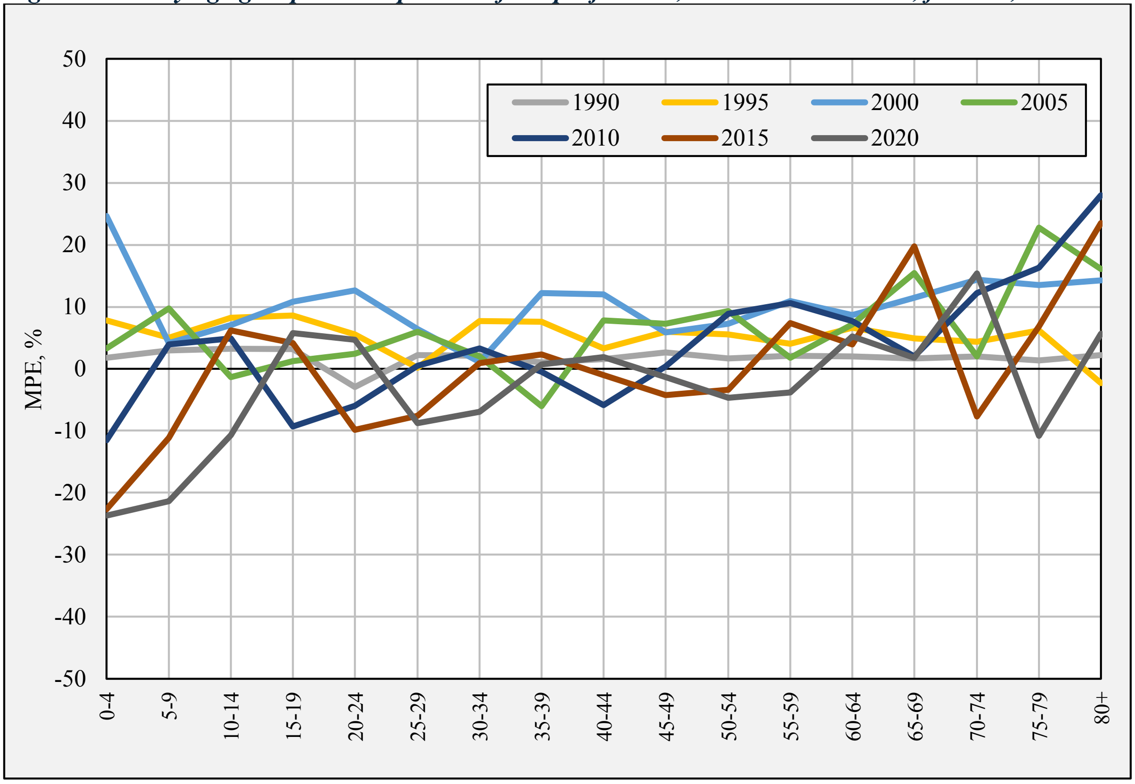
Source: The author’s calculations based on data from UN WPP 1994–2019 revisions.

Fig. 17: MPE by age groups across the duration of the projections, UN WPP 1994-2019, females, %



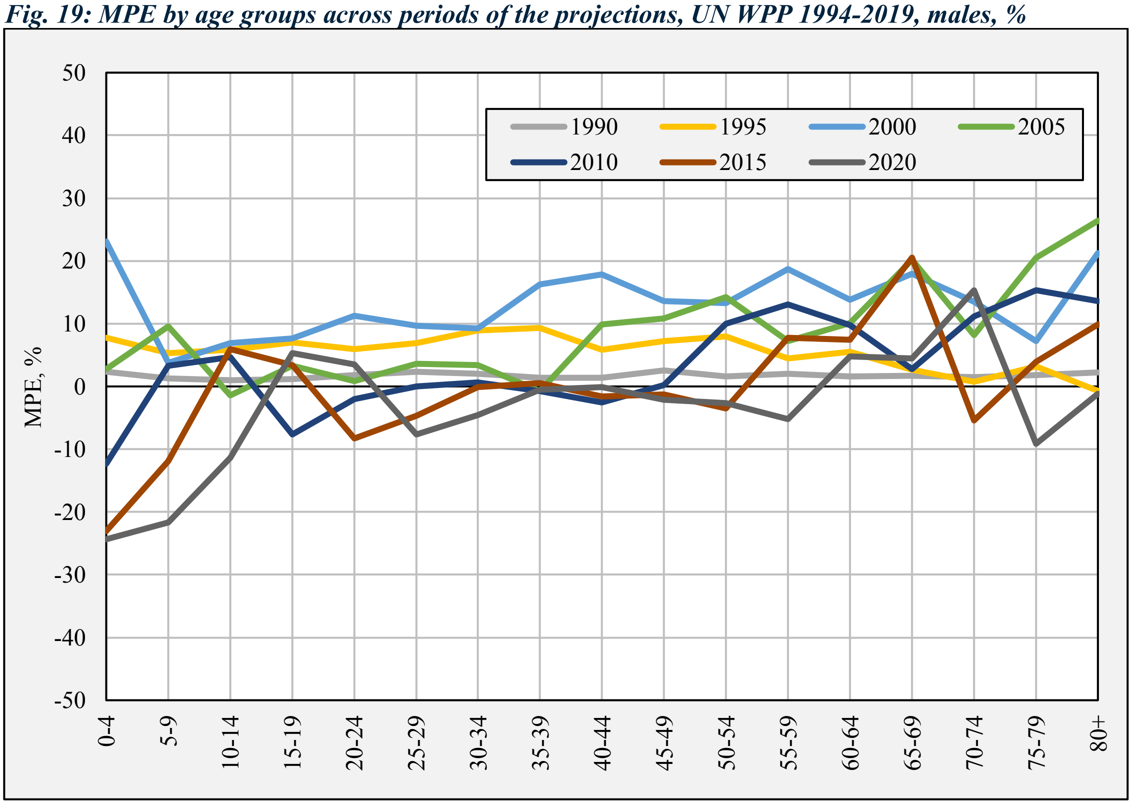
Source: The author’s calculations based on data from UN WPP 1994–2019 revisions.

Fig. 18: MPE by age groups across periods of the projections, UN WPP 1994-2019, females, %



Source: The author's calculations based on data from UN WPP 1994–2019 revisions.

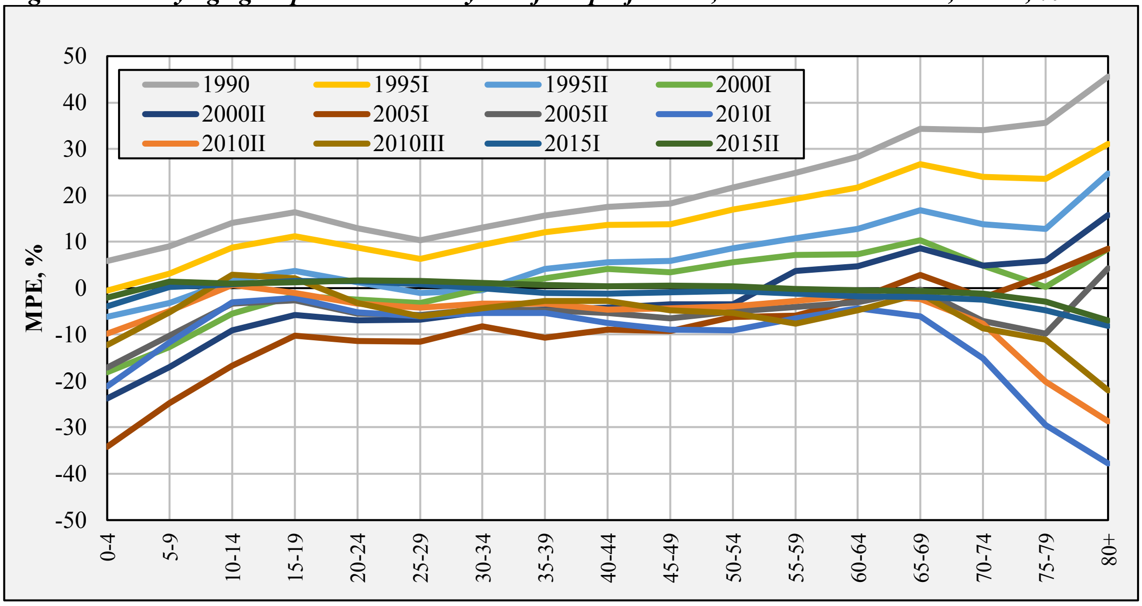
Fig. 19: MPE by age groups across periods of the projections, UN WPP 1994-2019, males, %



Source: The author's calculations based on data from UN WPP 1994–2019 revisions.

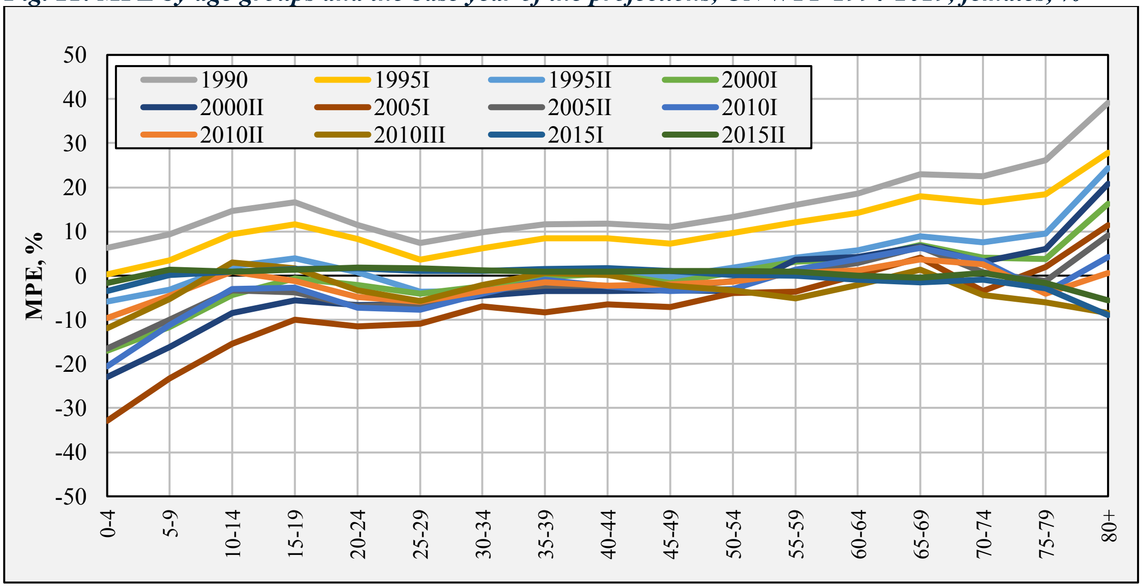
The above figures represent MPE across periods of the projections, UN WPP 1994–2019, females and males. Analysis reveals that during earlier projection periods (1990, 1995, 2000, and particularly 2005), the MPEs are predominantly positive across all age structures for both genders. This suggests an overestimation of population sizes in these projections. Conversely, in later projection periods (2010, 2015, and 2020), the projections tend to underestimate the population size in younger age groups while overestimating the population size in older age groups. Additionally, a recurring pattern is observed where an error noted in one age group for a given five-year period appears to propagate to the adjacent older age group in the subsequent five-year period. This cyclical error likely results from persistent inaccuracies in the assumptions related to age-specific demographic trends.

Fig. 20: MPE by age groups and the base year of the projections, UN WPP 1994-2019, males, %



Source: The author's calculations based on data from UN WPP 1994–2019 revisions.

Fig. 21: MPE by age groups and the base year of the projections, UN WPP 1994-2019, females, %



Source: The author's calculations based on data from UN WPP 1994–2019 revisions.

The above figures represent MPE by the base year of the population, calculated over the periods) for the UN WPP 1994-2019 revisions, for both females and males. The figures indicate that the initial two projections exhibited positive MPEs across all age groups, suggesting an overestimation of the population size at every age. In the case of males, projections with the jump-off year 2010 demonstrate an underestimation in older age groups, a reversal from earlier projections which overestimated the population in these groups. Furthermore, both male and female projections consistently underestimated the population in the youngest age groups.

Chapter 6

National Statistics

6.1 Fertility

To assess the accuracy of the UN's values compared to Kazakhstan's national statistics, which are traditionally calculated from January 1st to December 31st annually, values provided by the national statistics were recalculated to ensure comparability and accuracy when forecasting demographic trends for Kazakhstan. The following procedure was followed:

1. Initially, births and deaths were calculated for two distinct periods: the first six months (from January 1st to June 30th) and the last six months (from July 1st to December 31st). This ensures that the time periods are centered around the mid-year. Each observation begins on July 1st of a given year and concludes at mid-year five years later.
2. Subsequently, death rates were determined by dividing the number of deaths by the mid-year population of the respective period. Following this, life expectancy was calculated using the DERAS software.
3. A similar methodology was applied to calculate ASFR, which were then used to determine the observed TFR.

The analysis of errors between different revisions of the WPP and the National Statistics of Kazakhstan reveals the gradual improvement in the accuracy of TFR projections over time. The RMSE and the MAPE indicate significant errors across earlier revisions, with the highest errors in the UN 2004 revision (RMSE -0,3, MAPE 29,87%) and the lowest in the UN 1994 revision (RMSE -0,05, MAPE 16,33%). The MPE values reflect consistent underestimations, peaking in the UN 2004 revision (30,12%). However, more recent revisions, such as UN 2017, show marked improvements in accuracy, with MAPE and MPE at 9,08%, suggesting that methodological enhancements have led to projections that better align with actual national statistics.

Tab. 15: Errors in TFR projections between 1994–2019 revisions and Kazakhstan's National Statistics

	RMSE	MAPE (%)	MPE (%)
UN 1994	0,18	16,33	18,24
UN 1996	0,19	17,27	18,59
UN 1998	0,19	17,27	18,59
UN 2000	0,25	21,66	24,60

	RMSE	MAPE (%)	MPE (%)
UN 2002	0,26	23,08	26,21
UN 2004	0,30	29,87	30,12
UN 2006	0,18	16,35	18,00
UN 2008	0,21	21,07	21,50
UN 2010	0,11	10,11	10,85
UN 2012	0,13	12,11	12,88
UN 2015	0,10	10,49	10,49
UN 2017	0,09	9,08	9,08

Source: The Author's calculation based on data from UN WPP 1994–2019 statistics and Bureau of National Statistics

The table below presents the errors in the number of births projections for Kazakhstan. The MPE indicates a consistent underestimation of births across all revisions, with the highest underestimation in the UN 2008 revision (-26,71%), and the lowest in the UN 1998 revision (-6,13%). The MAPE shows the highest deviations in the UN 2002 revision (28,03%), and the lowest in the UN 2017 revision (8,25%), reflecting improved accuracy over time. The RMSE highlights significant discrepancies in the UN 2004 revision (0,84), and the most accurate projections in the UN 2017 revision (0,14).

Tab. 16: Errors in the Number of Births projections between 1998–2019 revisions and Kazakhstan National Statistics

	MPE (%)	MAPE (%)	RMSE
UN 1998	-6,13	22,92	0,50
UN 2000	-22,56	26,31	0,67
UN 2002	-26,40	28,03	0,75
UN 2004	-17,54	21,18	0,84
UN 2006	-20,98	20,98	0,55
UN 2008	-26,71	26,71	0,61
UN 2010	-14,85	14,85	0,27
UN 2012	-16,47	16,47	0,33
UN 2015	-9,71	9,71	0,16
UN 2017	-8,25	8,25	0,14

Source: The Author's calculation based on data from UN WPP 1998–2019 statistics and Bureau of National Statistics

6.2 Mortality

Although more recent revisions usually show a significant improvement, it can be seen from the tables above that considerable deviations still persist, stressing the challenges in accurately projecting mortality trends. Analyzing the errors in the number of deaths projections for Kazakhstan, it can be found that for males, the highest errors were observed in the UN 2012 revision, with the MPE of 29,02%, the MAPE of 29,02%, and the RMSE of 0,29, while the lowest errors were in the UN 2002 revision. For females, the largest errors were in the UN 2008 revision, with an MPE and MAPE of 22,75% and an RMSE of 0,23, while the lowest errors were in the UN 1998 revision.

Tab. 17: Errors in the Number of Deaths projections between 1998–2019 revisions and Kazakhstan National Statistics, males

	MPE (%)	MAPE (%)	RMSE
UN 1998	-9,67	11,78	0,13
UN 2000	6,02	6,44	0,08
UN 2002	-2,76	3,85	0,05
UN 2004	15,03	15,03	0,17
UN 2006	4,91	5,13	0,06
UN 2008	27,38	27,38	0,27
UN 2010	16,34	16,34	0,17
UN 2012	29,02	29,02	0,29
UN 2015	24,18	24,18	0,24
UN 2017	23,26	23,26	0,23

Source: The Author's calculation based on data from UN WPP 1998–2019 statistics and Bureau of National Statistics

Tab. 18: Errors in the Number of Deaths projections between 1998–2019 revisions and Kazakhstan National Statistics, females

	MPE (%)	MAPE (%)	RMSE
UN 1998	-0,09	7,08	0,08
UN 2000	10,38	10,38	0,11
UN 2002	3,66	5,85	0,07
UN 2004	16,17	16,17	0,17
UN 2006	9,24	9,24	0,10
UN 2008	22,75	22,75	0,23
UN 2010	16,31	16,31	0,16
UN 2012	16,26	16,26	0,17
UN 2015	18,57	18,57	0,19
UN 2017	17,23	17,23	0,17

Source: The Author's calculation based on data from UN WPP 1998–2019 statistics and Bureau of National Statistics.

The table below presents errors in life expectancy projections for males and females in Kazakhstan. The MPE indicates that earlier revisions, such as UN 1994, had significant overestimations, especially for males – 8,95%. At the same time the highest underestimations occurred in the UN 2004 revision for females (-5,17%) and in the UN 2008 revision for males (-8,80%). The MAPE shows the highest deviations in the UN 2004 revision for females (5,17%) and in the UN 1994 revision for males (8,95%), reflecting substantial errors in early projections. However, more recent revisions, particularly from 2000 onwards, demonstrate improved accuracy, with the lowest MAPE values observed in the UN 2002 revision for both females (1,11%) and males (1,35%).

Tab. 19: Errors in Life Expectancy between 1994–2019 revisions and Kazakhstan National Statistics, males and females

	MPE		MAPE	
	female	male	female	male
UN 1994	4,40	8,95	4,40	8,95
UN 1996	1,72	5,10	1,97	5,10
UN 1998	1,73	5,10	1,96	5,10
UN 2000	-1,91	-2,08	1,91	2,08
UN 2002	-0,38	-0,11	1,11	1,35
UN 2004	-5,17	-7,83	5,17	7,83
UN 2006	-1,05	-2,03	1,05	2,03
UN 2008	-4,13	-8,80	4,13	8,80
UN 2010	-3,19	-5,63	3,19	5,63
UN 2012	-4,17	-8,26	4,17	8,26
UN 2015	-2,80	-4,92	2,80	4,92
UN 2017	-2,38	-4,45	2,38	4,45

Source: The Author's calculation based on data from UN WPP 1998–2019 statistics and Bureau of National Statistics.

These discrepancies can be explained by the fact that, as mentioned earlier, in Kazakhstan during the 1990s and early 2000s, there was a high number of deaths due to diseases of the circulatory system. This led to an overestimation of life expectancy in the earlier revisions. Meanwhile, the revisions released in the mid-2000s maintained a scenario with high mortality, which subsequently resulted in an underestimation of life expectancy.

6.3 Migration

Tab. 20: Projected and Estimated values for Net Number of Migrants from UN WPP 1998–2019 and data from National Statistics, 1990–2020 (thousands)

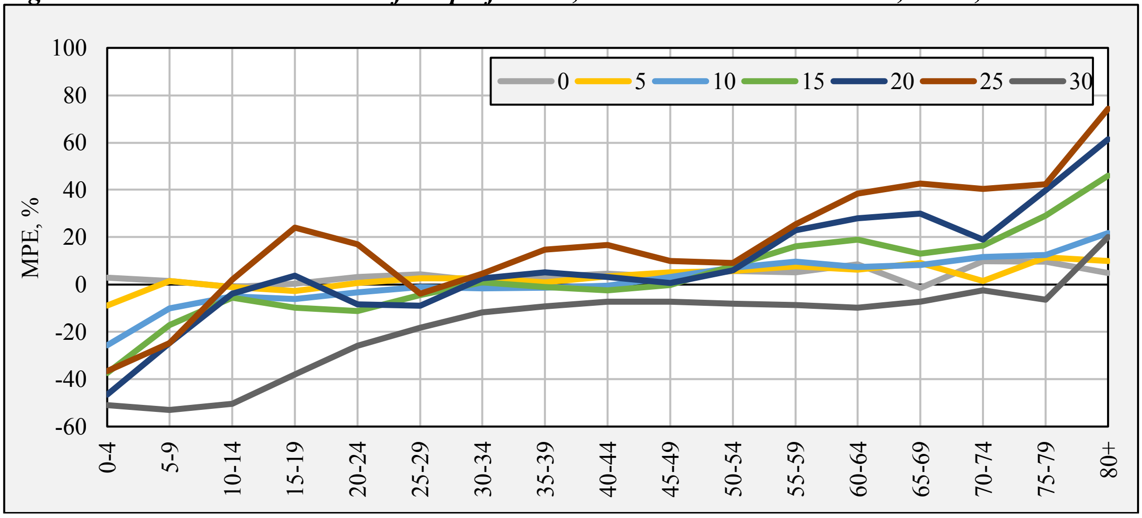
	1990–1995	1995–2000	2000–2005	2005–2010	2010–2015	2015–2020
UN 1998	-1200	-1088	-800	-400	-300	-200
UN 2000	-1000	-1000	-800	-600	-350	-300
UN 2002	-1164	-1488	-800	-743	-300	-300
UN 2004	-1509	-1320	-600	-400	-300	-300
UN 2006	-1509	-1320	-200	-200	-200	-200
UN 2008	-1509	-1320	-200	-100	-100	-100
UN 2010	-1509	-1320	-221	7	-50	-50
UN 2012	-1512	-1325	45	-11	0	0
UN 2015	-1512	-1325	45	-35	160	0
UN 2017	-1412	-1325	45	-35	160	0
KZ	***-1406***	***-1317***	***-199***	***72***	***-8***	***-121***

Sources: UN WPP 1998–2019 and Bureau of National Statistics.
Notes: Projected values are shaded. Data from National Statistics is in italics

From the table above, it can be seen that for the period 1995–2000, net migration was significantly higher compared to the projections presented in the UN WPP 1998 revision. The projection underestimated the Net Number of Migrants by 17,4%. For the period 2000–2005, the projections continued to show much higher negative figures (-800 000), while the observed value was -190 000 . In the 2005 to 2010 period, the observed data showed a positive net migration of 72 000, which was not reflected in the projections where it continued to show negative values up to -400 000. The periods 2010–2015 and 2015–2020 also saw significant differences, as the observed values turned negative at -8 000 and -121 000, respectively, while UN projections ranged from 0 to -300 000, indicating a persistent misalignment with actual migration trends.

6.4 Total Population

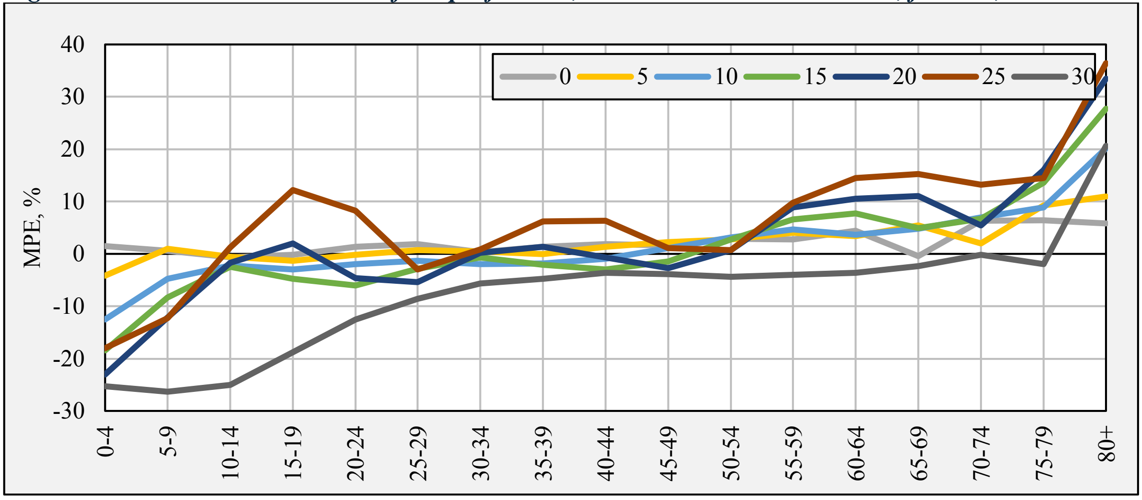
Fig. 22: MPE across the duration of the projections, UN WPP 1994-2019 and KZ, males, %



Source: The Author's calculation based on data from UN WPP 1994–2019 statistics and Bureau of National Statistics

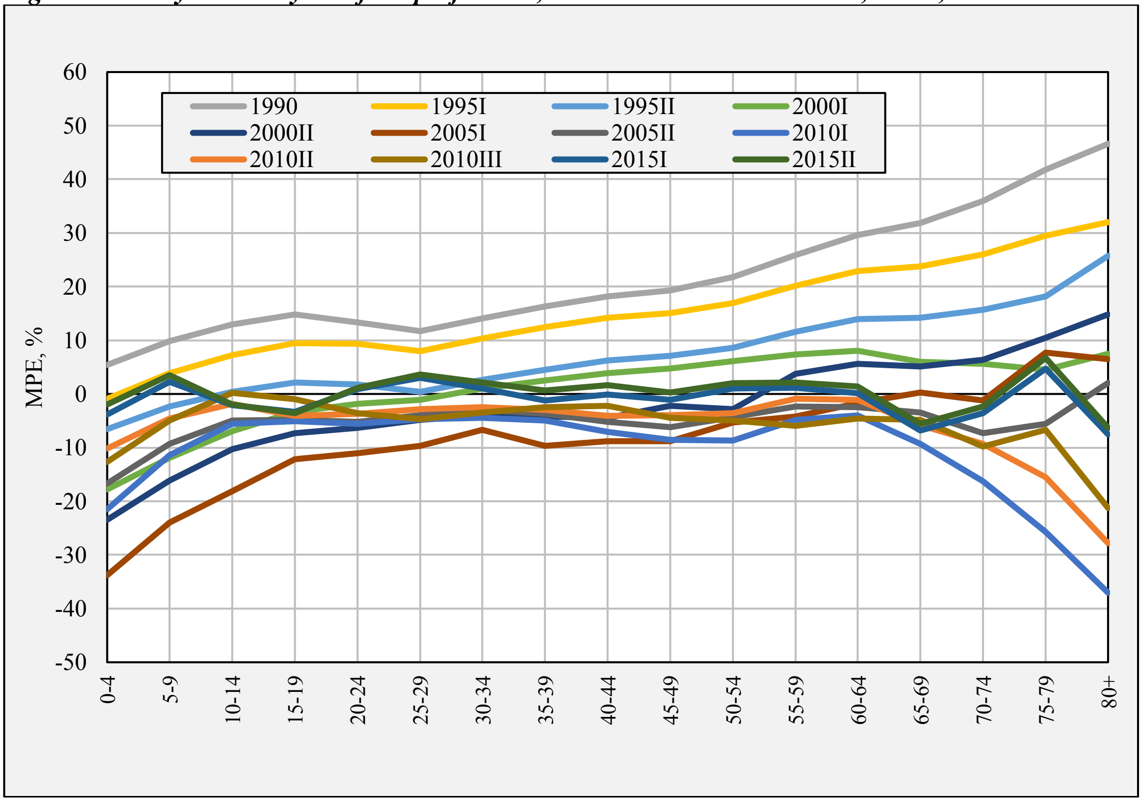
Analyzing the MPE by the base year of the projections reveals a trend where earlier projections (1990, 1995I) consistently overestimated the age-sex structure of the population. In contrast, later projections exhibit a tendency to underestimate this structure. Across all projections, there is a consistent pattern of underestimation in younger age groups and overestimation in older age groups, particularly pronounced in the male population.

Fig. 23: MPE across the duration of the projections, UN WPP 1994-2019 and KZ, females, %



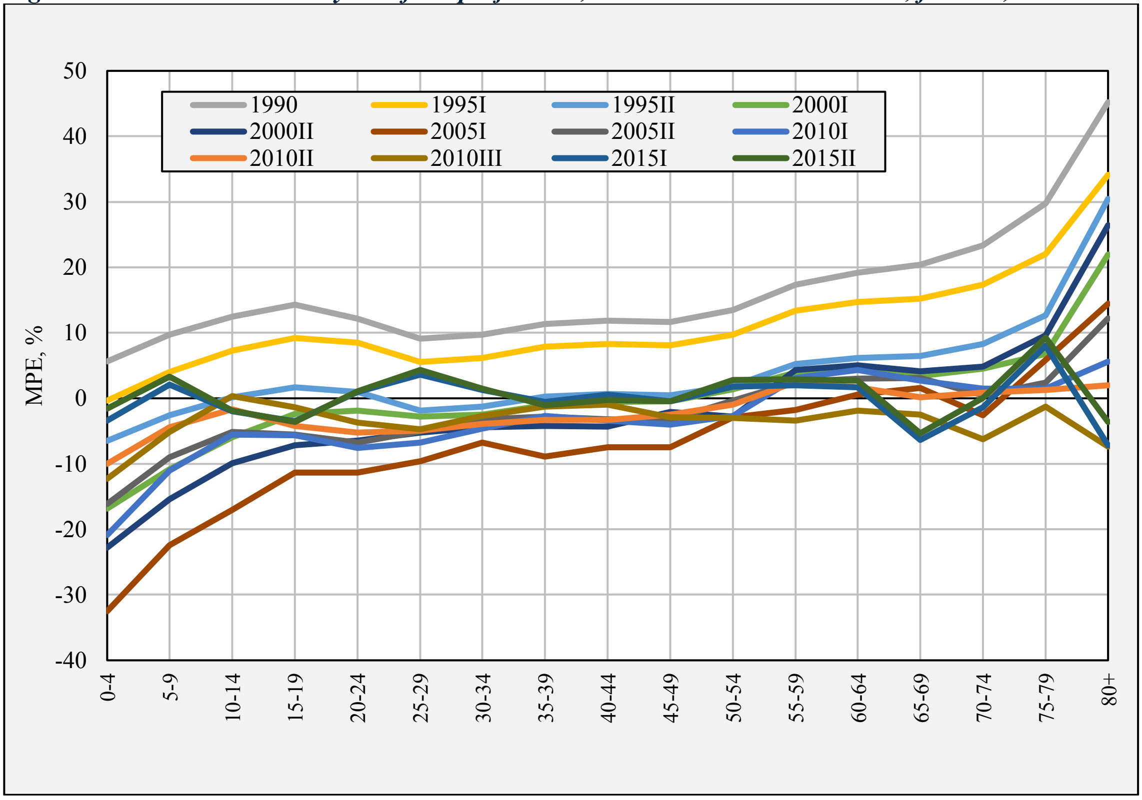
Source: The Author's calculation based on data from UN WPP 1994–2019 statistics and Bureau of National Statistics

Fig. 24: MPE by the base year of the projections, UN WPP 1994-2019 and KZ, males, %



Source: The Author's calculation based on data from UN WPP 1994–2019 statistics and Bureau of National Statistics

Fig. 25: MPE across the base year of the projections, UN WPP 1994-2019 and KZ, females, %

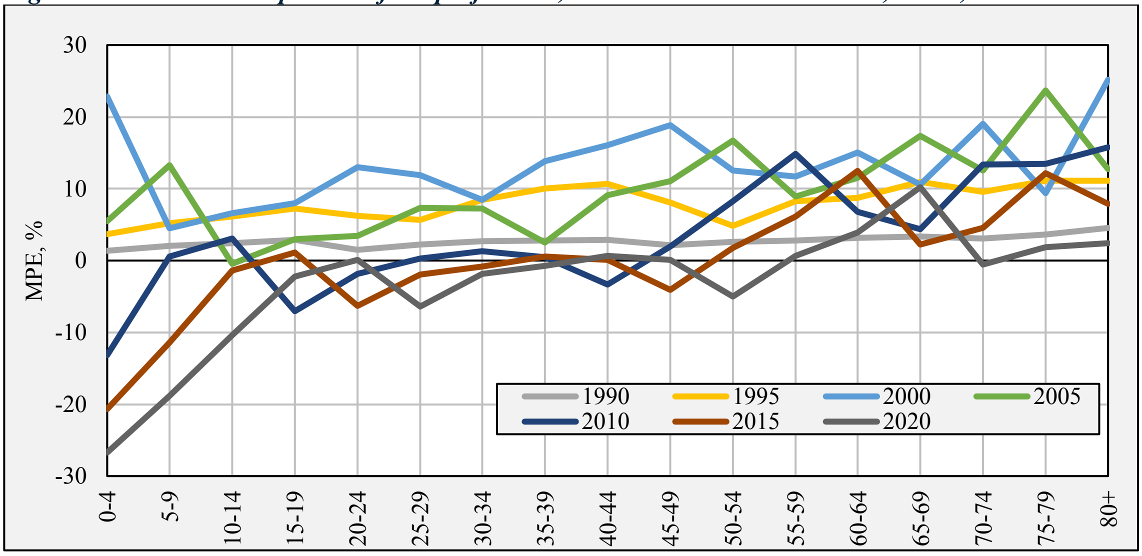


Source: The Author's calculation based on data from UN WPP 1994–2019 statistics and Bureau of National Statistics

For younger male age groups, such as 0–4 and 5–9, there were significant overestimations in 2000, with MPEs of 22,88% and 4,49%, respectively (Fig.25). These projections shifted to notable underestimations by 2020, highlighting difficulties in accurately forecasting birth rates and infant survival. Middle age groups, such as 10–14 and 15–19, showed fluctuations, with initial minor overestimations transitioning to underestimations in recent years, reflecting demographic changes or migration influences. Older age groups, from 20–24 to 80+, displayed a mix of errors, with significant overestimations in 2000, especially for ages 45-49 and 60-64, and some underestimations by 2020.

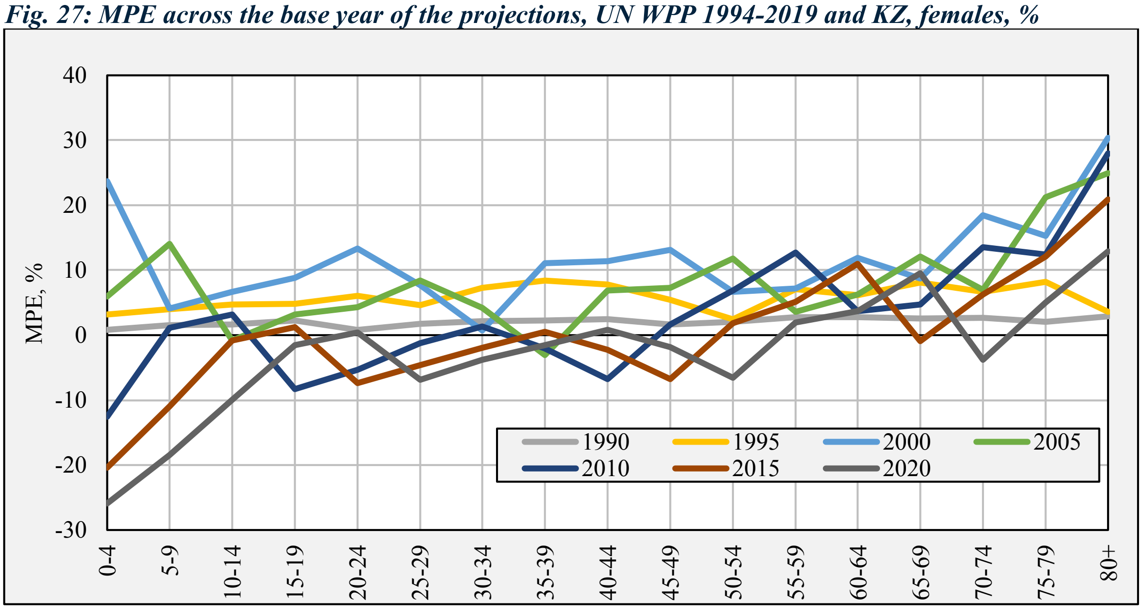
The MPE data for females shows that the youngest age group, 0–4, experienced significant overestimations in 2000 (23,73%), shifting to notable underestimations by 2020 (-25,91%). One of the reasons for the underestimation could be the increase in birth rates. The 5–9 age group showed slight overestimations in 2000 (4,15%) but by 2020 faced considerable underestimations (-18,44%). Meanwhile, the 45–49 to 60–64 age groups generally showed overestimations initially, transitioning to slightly underestimated values by 2020.

Fig. 26: MPE across the periods of the projections, UN WPP 1994-2019 and KZ, males, %



Source: The Author’s calculation based on data from UN WPP 1994–2019 statistics and Bureau of National Statistics

Fig. 27: MPE across the base year of the projections, UN WPP 1994-2019 and KZ, females, %



Source: The Author’s calculation based on data from UN WPP 1994–2019 statistics and Bureau of National Statistics

Conclusions

The adoption of efficient management decisions at different levels should be established on a reliable information base, an important part of which are the results of demographic forecasting. Despite the critical role of demographic forecasting in planning and policy-making, significant differences often exist between forecasted demographic data and actual outcomes.

While the methodology used by UN has evolved and improved over time, the accuracy of projections is greatly influenced by the quality and relevance of the underlying data, as well as the assumptions regarding socio-economic and demographic trends.

This thesis has examined the reliability of the UN DESA demographic projections for Kazakhstan, covering the period from 1994 to 2019. Through a systematic analysis of errors across different age and sex groups, this study has identified significant differences between projected and actual population figures, revealing both overestimations and underestimations at various times.

Key findings indicate that short-term projections tend to be more accurate than long-term projections, consistent with global trends in demographic forecasting. Notably, the projections have generally overestimated the population size in older age groups while underestimating younger age groups. This pattern of error suggests issues with the assumptions used about fertility and mortality rates, which may not have adequately captured recent changes in demographic behavior or been adjusted based on updated data.

The analysis has highlighted the importance of incorporating high-quality and timely data into the forecasting process. Kazakhstan's demographic policies, particularly those aimed at increasing birth rates and managing migration, have had a significant impact on demographic trends, yet these influences were not always reflected accurately in the projections. The study underscores the need for forecasters to adapt their models to reflect the changing socio-economic conditions and policy environments within the country. It was evident that UN experts were not always aware of the unique reproductive behavior of the Kazakhstani population, nor could they always discern if certain behavioral changes (such as increased mortality) were temporary phenomena. This was apparent in the scenarios chosen by the experts. As observed, users often interpret projection results as projections, indicating likely future developments based on the forecaster's current understanding. Therefore, it is imperative for agencies like UN DESA to clearly communicate the assumptions underlying their projections to ensure that users have a correct understanding of what the projections might entail.

The methodology employed by UN DESA, despite its sophistication, did not consistently translate into accurate forecasts. The projections' accuracy appeared less influenced by the chosen methodology and more by the quality and recency of the input data. This suggests that while advanced statistical models and methods are crucial, their effectiveness is highly dependent on the timeliness and relevance of the data inputs. For instance, the significant errors observed in projections from the 1990s and early 2000s can be partly attributed to the UN's failure to fully incorporate the effects of Kazakhstan's economic transition post-independence, which had significant impacts on fertility, mortality, and migration patterns

Ultimately, it is advised that the country prepare its own population forecasts, as local experts will take into account the specific characteristics of Kazakhstan's population.

REFERENCES

AHLBURG, D. A., LUTZ, W., & VAUPEL, J. W. (1998). Ways to Improve Population Forecasting: What Should Be Done Differently in the Future? Population and Development Review, 24, 191–198. https://doi.org/10.2307/2808056

ALEKSEENKO A.N., AUBAKIROVA J.S, and ZHANBOSINOVA A.S. (2019). Echotomographic evolutions and the formation of a sovereign demographic system in Kazakhstan https://doi.org/10.21638/11701/spbu02.2019.413

ALIMBEKOVA G.T., ZEYNALOVA E.A. (n.d). Migration in Kazakhstan in the post-Soviet period. https://ciom.kz/upload/userfiles/files/Migratsiya_postsovetski_period.pdf

ALKEMA, L., RAFTERY, A. E., GERLAND, P., CLARK, S. J., PELLETIER, F., BUETTNER, T. AND HEILIG, G. K. (2011). Probabilistic projections of the total fertility rate for all countries. Demography 48 815–839.

BEKHOZHAEVA A.K. (2004). Regional trends in the demographic development of the Republic of Kazakhstan in 1990-2002. *Questions of Statistics* 11: 69-73.

BUREAU OF NATIONAL STATISTICS. (2022).

DEMOGRAPHIC AND HEALTH SURVEY. (1995). Demographic and Health Survey: Kazakstan.

FILDES, R., & GOODWIN, P. (2007). Against Your Better Judgment? How Organizations Can Improve Their Use of Management Judgment in Forecasting. *Interfaces*, *37*(6), 570–576. http://www.jstor.org/stable/20141547

CHRISTENSEN, P., GILLINGHAM, K., & NORDHAUS, W. (2018). Uncertainty in forecasts of long-run economic growth. Proceedings of the National Academy of Sciences of the United States of America, 115(21), 5409–5414. https://doi.org/10.1073/pnas.1713628115

GERLAND, P., RAFTERY, A. E., SEVČÍKOVÁ, H., LI, N., GU, D., SPOORENBERG, T., ALKEMA, L., FOSDICK, B. K., CHUNN, J., LALIC, N., BAY, G., BUETTNER, T., HEILIG, G. K., & WILMOTH, J. (2014). World population stabilization unlikely this century. *Science (New York, N.Y.)*, *346*(6206), 234–237. https://doi.org/10.1126/science.1257469

GOODWIN, P. AND LAWTON, R. (1999) On the Asymmetry of the Symmetric MAPE. *International Journal of Forecasting* 15: 405-408. http://dx.doi.org/10.1016/S0169-2070(99)00007-2

HYNDMAN, R. J., & KOEHLER, A. B. (2006). Another look at measures of forecast accuracy. International Journal of Forecasting 22: 679–688

KEILMAN, N. (1986). The unpredictability of population trends. Impact Assessment, 4, 49-80. doi:10.1080/07349165.1986.9725778

KEILMAN, N. (1990). *Uncertainty in National Population Forecasting: Issues, Measurement, Backgrounds, Recommendations: with an Application to the Netherlands*. Swets & Zeitlinger.

KEILMAN, N. (2001). Uncertain population forecasts. *Nature* 412:490–491. https://doi.org/10.1038/35087685

KEYFITZ, N. (1981). The Limits of Population Forecasting. *Population and Development Review*, *7*(4), 579–593. https://doi.org/10.2307/1972799

KOLASSA, S. (2016). Evaluating predictive count data distributions in retail sales forecasting. International Journal of Forecasting 32: 788–803

KOLASSA S., MARTIN R. (2011). Foresight: The International Journal of Applied Forecasting. *International Institute of Forecasters* 23: 21-27, Fall.

KOUTSANDREAS D., SPILIOTIS E., PETROPOULOS F., AND ASSIMAKOPOULOS V. (2022). On the selection of forecasting accuracy measures. Journal of the Operational Research Society 73 (5): 937-954. https://doi.org/10.1080/01605682.2021.1892464

LITVINETS Y.Y. (2006). Soviet Demographic Policy in the 1980s. https://elar.urfu.ru/bitstream/10995/23283/1/iurg-2006-47-36.pdf

LIU P., RAFTERY A.E. (2018). Accounting for Uncertainty About Past Values In Probabilistic Projections of the Total Fertility Rate for All Countries. University of Washington

LUTZ, W., KC, S. (2010). Dimensions of global population projections: what do we know about future population trends and structures? Philosophical Transactions of the Royal Society B: Biological Sciences, 365, 2779 - 2791.

MAKRIDAKIS S. (1993). Accuracy measures: theoretical and practical concerns. *International Journal of Forecasting* 9 (4): 527-529

MARTINE G., ALVES J.E., AND CAVENAGHI S. (2013). Urbanization and Fertility Decline: Cashing in on Structural Change. IIED Working Paper. IIED, London.

NATIONAL RESEARCH COUNCIL. (2000). Beyond Six Billion: Forecasting the World's Population. Washington, DC: The National Academies Press. doi: 10.17226/9828.

RAFTERY, A. E., CHUNN, J. L., GERLAND, P., & SEVČÍKOVÁ, H. (2013). Bayesian probabilistic projections of life expectancy for all countries. *Demography*, *50*(3), 777–801. https://doi.org/10.1007/s13524-012-0193-x

RYBALKIN V.S., VARNAVINA A.O. (2017). Migration processes in the Republic of Kazakhstan in the late XX - early XXI century. The post-Soviet continent, (2 (14)): 48-58.

TOFALLIS C. (2014). A Better Measure of Relative Prediction Accuracy for Model Selection and Model Estimation. *Journal of the Operational Research Society* (2015) 66: 1352–1362. https://ssrn.com/abstract=2635088

UNITED NATIONS DEVELOPMENT PROGRAM. (1995). Kazakstan human development report 1995. Almaty, Kazakstan: UNDP.

UNITED NATIONS (2019a). How certain are the United Nations global population projections? Population Facts. New York: Department of Economic and Social Affairs, Population Division.

WILSON, T. (2007). The forecast accuracy of australian bureau of statistics national population projections. Journal of Population Research, 24(1), 91–117. http://www.jstor.org/stable/41110880

UNITED NATIONS. *(1995).* World Population Prospects, The 1994 revision.

UNITED NATIONS. (1998). World Population Prospects: The 1996 revision.

UNITED NATIONS. (1999). World Population Prospects: The 1998 revision.

UNITED NATIONS. (2001). World Population Prospects: The 2000 revision.

UNITED NATIONS. (2002). World Population Prospects: The 2002 revision

UNITED NATIONS. (2005). World Population Prospects: The 2004 revision.

UNITED NATIONS. (2007). World Population Prospects: The 2006 revision.

UNITED NATIONS. (2010). World Population Prospects: The 2008 revision.

UNITED NATIONS. (2011). World Population Prospects: The 2010 revision.

UNITED NATIONS. (2013). World Population Prospects: The 2012 revision.

UNITED NATIONS. (2015). World Population Prospects: The 2015 revision.

UNITED NATIONS. (2017). World Population Prospects: The 2017 revision.

UNITED NATIONS. (2019). World Population Prospects: The 2019 revision.

UNITED NATIONS. (2022). World Population Prospects: The 2022 revision.

APPENDICES

List of appendices

Annex 1: MPE by age groups and the base year of projections, males

	Base year of UN WPP projection											
	1990	1995I	1995II	2000I	2000II	2005I	2005II	2010I	2010II	2010III	2015I	2015II
0-4	5,42	-0,91	-6,55	-17,80	-23,45	-33,78	-16,69	-21,43	-10,16	-12,60	-3,80	-1,94
5-9	9,83	3,90	-2,38	-11,87	-16,17	-23,95	-9,25	-11,47	-4,59	-4,91	2,25	3,51
10-14	12,90	7,29	0,44	-6,95	-10,35	-18,14	-4,99	-5,59	-1,85	0,21	-2,06	-2,02
15-19	14,85	9,43	2,12	-3,35	-7,35	-12,13	-4,86	-5,09	-4,07	-1,00	-3,27	-3,53
20-24	13,33	9,32	1,80	-1,83	-6,34	-11,01	-5,16	-5,55	-3,66	-3,63	0,91	1,09
25-29	11,71	7,93	0,43	-1,14	-4,80	-9,68	-3,94	-4,72	-2,79	-4,64	2,98	3,63
30-34	14,09	10,30	2,67	0,98	-3,99	-6,73	-2,92	-4,48	-2,47	-3,45	1,05	2,17
35-39	16,33	12,49	4,54	2,50	-4,24	-9,65	-3,74	-4,97	-2,93	-2,41	-1,18	0,63
40-44	18,20	14,21	6,19	3,86	-4,45	-8,81	-5,22	-7,05	-4,12	-2,22	-0,07	1,62
45-49	19,29	15,03	7,07	4,69	-2,17	-8,80	-6,20	-8,61	-4,00	-4,44	-1,13	0,24
50-54	21,79	16,91	8,59	6,17	-2,85	-5,29	-4,30	-8,70	-3,57	-5,00	1,03	1,97
55-59	25,84	20,20	11,55	7,41	3,79	-4,20	-2,38	-4,77	-0,89	-6,00	1,09	2,14
60-64	29,65	22,89	13,91	8,05	5,57	-1,66	-2,49	-4,00	-1,09	-4,54	0,18	1,42
65-69	31,81	23,83	14,19	5,97	5,13	0,22	-3,43	-9,31	-5,78	-4,81	-6,80	-5,60
70-74	36,01	26,02	15,66	5,60	6,41	-1,23	-7,29	-16,32	-9,27	-9,81	-3,63	-2,38
75-79	41,78	29,52	18,23	4,49	10,53	7,71	-5,53	-25,72	-15,51	-6,73	4,72	6,73
80+	46,62	32,02	25,74	7,52	14,83	6,52	2,05	-37,05	-27,81	-21,25	-7,52	-6,31

Source: UN WPP 1994–2019, Bureau of National Statistics of Kazakhstan.
Notes: UN projections were compared to Kazakhstan's official statistical data.

Annex 2:MPE by the base year of projections, females

	Base year of UN WPP projection											
	1990	1995I	1995II	2000I	2000II	2005I	2005II	2010I	2010II	2010III	2015I	2015II
0-4	5,64	-0,35	-6,49	-16,84	-22,80	-32,50	-16,12	-20,87	-9,99	-12,29	-3,39	-1,58
5-9	9,68	4,06	-2,58	-10,86	-15,45	-22,45	-8,95	-11,08	-4,39	-5,09	2,14	3,36
10-14	12,50	7,25	0,14	-5,96	-9,89	-17,00	-5,12	-5,55	-1,62	0,34	-1,95	-1,91
15-19	14,25	9,20	1,72	-2,35	-7,16	-11,37	-5,52	-5,69	-4,22	-1,41	-3,46	-3,63
20-24	12,20	8,51	1,00	-1,83	-6,41	-11,32	-6,75	-7,60	-5,20	-3,69	0,92	1,11
25-29	9,07	5,54	-1,86	-2,79	-5,21	-9,65	-5,00	-6,72	-4,92	-4,77	3,64	4,29
30-34	9,71	6,13	-1,30	-2,49	-4,41	-6,76	-3,16	-4,67	-3,95	-2,67	1,29	1,51
35-39	11,36	7,86	0,25	-1,05	-4,21	-8,93	-2,84	-2,73	-3,26	-1,26	-0,53	-1,08
40-44	11,88	8,35	0,72	-0,51	-4,29	-7,48	-3,22	-3,38	-3,35	-0,96	0,58	-0,25
45-49	11,66	8,04	0,48	-0,50	-2,12	-7,49	-3,31	-4,06	-2,52	-2,90	-0,50	-0,47
50-54	13,50	9,73	1,84	1,40	-2,79	-2,89	-0,33	-2,65	-0,97	-2,96	1,80	2,81
55-59	17,31	13,43	5,29	3,96	4,29	-1,78	2,51	3,12	2,49	-3,41	2,02	2,89
60-64	19,20	14,71	6,18	4,50	5,07	0,52	2,97	4,30	1,71	-1,83	1,73	2,66
65-69	20,44	15,26	6,45	3,46	4,11	1,57	3,07	2,67	0,16	-2,44	-6,32	-5,34
70-74	23,34	17,34	8,31	4,49	4,80	-2,59	0,48	1,51	0,90	-6,20	-1,41	0,02
75-79	29,72	21,99	12,63	6,80	9,67	5,85	2,36	1,46	1,27	-1,22	8,03	9,27
80+	45,20	34,12	30,44	21,95	26,43	14,45	12,21	5,61	2,02	-7,37	-7,18	-3,65

Source: UN WPP 1994–2019, Bureau of National Statistics of Kazakhstan.
Notes: UN projections were compared to Kazakhstan's official statistical data.

Annex 3: MPE by age groups across the periods, males

	Periods						
	1990	1995	2000	2005	2010	2015	2020
0-4	1,37	3,67	22,88	5,46	-13,17	-20,68	-26,72
5-9	2,08	5,23	4,49	13,25	0,61	-11,44	-18,82
10-14	2,39	6,13	6,55	-0,44	3,04	-1,37	-10,43
15-19	2,90	7,29	7,97	3,01	-7,01	1,12	-2,19
20-24	1,49	6,24	13,00	3,43	-1,82	-6,33	0,13
25-29	2,20	5,66	11,92	7,34	0,24	-1,96	-6,39
30-34	2,74	8,42	8,41	7,28	1,30	-0,79	-1,83
35-39	2,82	10,04	13,83	2,55	0,48	0,54	-0,72
40-44	2,86	10,66	16,05	9,07	-3,32	0,06	0,64
45-49	2,18	8,06	18,83	11,09	1,93	-4,11	0,10
50-54	2,57	4,80	12,53	16,74	8,24	1,76	-4,98
55-59	2,80	8,26	11,75	8,87	14,83	6,12	0,66
60-64	3,17	8,76	15,05	11,48	6,76	12,49	3,91
65-69	3,39	10,99	10,56	17,36	4,40	2,25	10,21
70-74	3,08	9,58	19,03	12,55	13,38	4,56	-0,56
75-79	3,64	11,11	9,39	23,68	13,44	12,13	1,86
80+	4,55	11,11	25,16	12,74	15,77	7,89	2,41

Source: UN WPP 1994–2019, Bureau of National Statistics of Kazakhstan.
Notes: UN projections were compared to Kazakhstan's official statistical data.

Annex 4: MPE by age groups across the periods, females

	Periods						
	1990	1995	2000	2005	2010	2015	2020
0-4	0,87	3,15	23,73	5,96	-12,56	-20,40	-25,91
5-9	1,53	3,97	4,15	14,07	1,13	-10,97	-18,44
10-14	1,65	4,74	6,66	-0,78	3,21	-0,83	-9,92
15-19	2,30	4,84	8,80	3,14	-8,27	1,23	-1,51
20-24	0,80	6,04	13,34	4,31	-5,33	-7,34	0,40
25-29	1,79	4,60	7,69	8,41	-1,25	-4,64	-6,85
30-34	2,15	7,24	0,73	4,24	1,32	-1,95	-3,78
35-39	2,22	8,40	11,13	-3,07	-2,02	0,54	-1,55
40-44	2,49	7,78	11,41	6,86	-6,77	-2,24	0,85
45-49	1,66	5,40	13,13	7,29	1,61	-6,72	-1,88
50-54	2,09	2,43	6,70	11,82	6,90	1,81	-6,51
55-59	2,75	7,03	7,22	3,60	12,75	5,12	2,00
60-64	2,79	6,16	11,93	6,23	3,68	11,00	3,69
65-69	2,54	8,08	8,77	12,15	4,69	-0,95	9,54
70-74	2,68	6,63	18,47	6,93	13,54	6,22	-3,77
75-79	2,05	8,18	15,32	21,25	12,44	12,13	5,07
80+	2,90	3,56	30,38	24,94	27,99	20,89	12,90

Source: UN WPP 1994–2019, Bureau of National Statistics of Kazakhstan.
Notes: UN projections were compared to Kazakhstan's official statistical data.

Annex 5: MPE by age groups and duration of a projection, males

	Duration of a projection						
	0	5	10	15	20	25	30
0-4	1,26	-4,70	-13,12	-18,94	-23,46	-18,54	-25,54
5-9	0,86	0,55	-5,31	-8,83	-12,51	-12,54	-26,64
10-14	-0,43	-0,46	-2,64	-3,09	-2,07	0,89	-25,40
15-19	0,58	-1,28	-3,04	-5,04	1,73	11,98	-19,22
20-24	1,75	0,72	-1,45	-5,11	-3,71	8,71	-13,20
25-29	2,39	1,94	0,28	-1,51	-3,37	-0,90	-9,47
30-34	1,26	2,15	0,53	1,48	2,33	3,76	-6,06
35-39	1,70	1,02	0,53	1,00	3,85	8,45	-4,53
40-44	2,57	2,06	0,43	0,57	3,75	10,43	-3,62
45-49	1,76	3,04	2,07	1,26	3,32	8,81	-3,27
50-54	2,98	2,84	4,00	5,30	5,33	8,47	-3,75
55-59	2,33	3,76	4,92	9,50	13,93	15,68	-4,69
60-64	4,00	2,93	3,69	11,17	17,52	23,97	-6,06
65-69	-1,01	3,56	3,58	8,04	18,84	27,53	-4,94
70-74	3,47	-0,53	4,71	9,72	13,63	27,14	-2,10
75-79	3,26	2,25	3,70	15,63	23,82	27,93	-4,48
80+	-0,83	-0,90	1,61	18,27	27,93	37,87	-0,53

Source: UN WPP 1994–2019, Bureau of National Statistics of Kazakhstan.
Notes: UN projections were compared to Kazakhstan's official statistical data.

Annex 6: MPE by age groups and duration of a projection, females

	Duration of a projection						
	0	5	10	15	20	25	30
0-4	1,53	-4,16	-12,53	-18,43	-23,04	-18,01	-25,29
5-9	0,61	0,91	-4,71	-8,33	-12,15	-12,23	-26,33
10-14	-0,66	-0,59	-2,21	-2,53	-1,76	1,19	-25,05
15-19	-0,23	-1,35	-3,02	-4,81	1,97	12,18	-18,81
20-24	1,31	-0,16	-1,95	-6,09	-4,60	8,24	-12,52
25-29	1,87	0,71	-1,34	-2,85	-5,42	-3,02	-8,63
30-34	0,36	0,41	-1,99	-0,64	0,25	0,84	-5,65
35-39	1,38	-0,02	-1,88	-2,03	1,34	6,18	-4,73
40-44	1,90	1,40	-0,93	-2,93	-0,66	6,31	-3,60
45-49	1,58	2,24	1,14	-1,48	-2,69	1,06	-3,90
50-54	2,83	2,77	3,16	2,81	0,72	0,66	-4,43
55-59	2,76	4,08	4,72	6,63	8,91	9,77	-3,99
60-64	4,44	3,44	3,63	7,73	10,53	14,54	-3,62
65-69	-0,42	5,42	4,79	4,96	11,03	15,24	-2,38
70-74	6,39	1,98	7,00	6,72	5,44	13,21	-0,22
75-79	6,41	9,24	8,83	13,53	15,97	14,47	-1,91
80+	5,86	10,96	20,15	27,72	33,46	36,42	20,63

Source: UN WPP 1994–2019, Bureau of National Statistics of Kazakhstan.
Notes: UN projections were compared to Kazakhstan's official statistical data.

Annex 7: Kazakhstan census data (1959, 1970, 1979, 1989, 1999, 2009).

	Census					
	1959	1970	1979	1989	1999	2009
Total Population, thousands	9283,2	13013,6	14688,3	16199,2	14953,1	16009,6
Population growth, thousands		3730,3	1674,8	1510,8	-1246,0	1056,5
Urban	4079,3	6745,4	7912,5	9182,6	8377,3	8662,4
Rural	5203,9	6457,5	6775,8	7016,5	6575,8	7347,2
Urban, %	43,9	51,8	53,9	56,7	56,0	54,1
Rural, %	56,1	48,2	46,1	43,3	44,0	45,9

Source: UN WPP 1994–2019, Bureau of National Statistics of Kazakhstan.