

WORKING PAPER

INAPP WP n. 23

Annuities in (N)DC Pension Schemes: Design, Heterogeneity, and Estimation Issues

Edward Palmer

Yuwei Zhao de Gosson de Varennes

Annuities in (N)DC Pension Schemes: Design, Heterogeneity, and Estimation Issues

Edward Palmer

edward.e.palmer@gmail.com

Yuwei Zhao de Gosson de Varennes

yuweidegossondevarennes@gmail.com

LUGLIO 2019

CONTENTS: 1. Introduction. – 2. The NDC annuity; 2.1 The basic annuity formula; 2.2 Policy considerations regarding the rate of return to be included in the annuity calculation; 2.3 The payment profile over time: A distributional issue; 2.4 The risk for the insurer; 2.5 The role of life expectancy in creating affordable, sustainable NDC pensions. – 3. Projecting life expectancy; 3.1 Different approaches to projecting life expectancy; 3.1.1 The period method for projecting life expectancy; 3.1.2 The statistical approach for cohort modeling – the Lee-Carter model and its variants; 3.1.3 The data analytical period-cohort approach: The Palmer-Alho-de Gosson (PAD) model; 3.2 A comparison of the period model, the LC model, and the PAD model for contemporary birth cohorts; 3.3 Policy implications. – 4. Variable annuities: What is gained and lost by recalculating the annuity at regular intervals based on new projections after the “normal” retirement age?. – 5. Socioeconomic determinants of life expectancy in the context of creating the annuity; 5.1 Higher life expectancy accompanies higher income, higher education, and choice of occupation; 5.2 Gender differences in income, life expectancy, and NDC annuity construction; 5.3 Other approaches to dealing with the socioeconomic gap in life expectancy in creating (N)DC annuities; 5.4 To what extent does unisex pooling level the gender playing field?. – 6. Conclusions; 6.1 Construction of the annuity and options for inclusion of the rate of return; 6.2 Unbiased estimation of cohort life expectancy; 6.3 A variable annuity?; 6.4 Socioeconomic heterogeneity in life expectancy?. – References. – Appendix. A formal description of the PAD model

INAPP – Istituto nazionale per l’analisi delle politiche pubbliche

Corso d’Italia 33
00198 Roma, Italia

Tel. +39 06854471
Email: urp@inapp.org

www.inapp.org

ABSTRACT

Annuities in (N)DC Pension Schemes: Design, Heterogeneity, and Estimation Issues*

This paper identifies and discusses four issues in creating annuities in (nonfinancial) defined contribution (NDC) schemes that are essential for systems' financial stability and fair inter/intragenerational redistribution. The first issue is the choice between incorporating the rate of return into the annuity or into the exogenous indexation. The second issue is in choosing a projection method for life expectancy that produces systematically unbiased estimates. The third issue is at what age the projection of life expectancy is to be fixed over the remaining lifetime of the annuity. The final issue is the prevalence of socioeconomic heterogeneity within the insurance pool.

KEYWORDS: Life Expectancy, Annuity, Defined Contribution Pension Schemes

JEL CODES: H55, J11

Abbreviations and Acronyms

DC	Defined Contribution
FDC	Financial Defined Contribution
LC	Lee-Carter
NDC	Nonfinancial Defined Contribution
PAD	Palmer-Alho-de Gosson

* ACKNOWLEDGMENTS - This paper is written for *Progress and Challenges of Nonfinancial Defined Contribution Pension Schemes, Volume 1: Addressing Marginalization, Polarization, and the Labor Market*, edited by Robert Holzmann, Edward Palmer, Robert Palacios, and Stefano Sacchi, to be published by the World Bank in autumn 2019. The authors are grateful to Robert Holzmann and Ronald Lee for comments and suggestions in various phases of our work and to Amy Gautam for first-rate copy editing. A first version of the paper was presented at the NDC III conference in Rome, October 5-6, 2017, and we are thankful to the participants for their comments and encouragement. The views expressed herein are those of the authors and do not necessarily reflect the views of the institutions they are associated with or the views of the World Bank and of the Inapp.

1. Introduction

When they emerged in the mid-1990s, nonfinancial defined contribution (NDC) schemes constituted a new way of thinking about the construction of the benefit in a universal public pension scheme. Individuals pay contributions to the public NDC scheme that are registered on personal accounts that receive a yearly rate of return. The rate of return is the rate of growth of the wage base upon which contributions are based. Contributions are paid to the scheme and noted on the individual's personal account throughout a lifetime working career until the individual claims a benefit after reaching the minimum pension age. The benefit received – in technical terms, the life annuity – is calculated as the individual's account balance divided by a divisor that depends on the average life expectancy of the individual's birth cohort – and possibly even a long-run assumed rate of return (i.e., a discount rate).

This paper presents and discusses policy issues that arise in the formulation of the annuity. Four aspects of annuity creation are examined in four separate sections of this paper. The topics are:

Construction of the annuity. It is a given that the annuity includes projected life expectancy at the annuitant's chosen age of retirement. There are options that need to be considered in how the rate of return enters into the picture, however. The choice is between (i) including the rate of return endogenously together with life expectancy in the construction of the annuity, (ii) basing the annuity solely on life expectancy complemented with exogenous yearly indexation, or (iii) employing a mixed version. What are the micro and macro implications for the distribution of pension pool payments over the life of a cohort of pensioners in the pension pool?

The importance of getting the projection of life expectancy right. Fulfillment of the NDC conditions of financial sustainability and intergenerational fairness require that the value of life expectancy employed in the formulation of the annuity does not systematically over- or underestimate life expectancy, viewed over a continuous series of birth cohorts. This section

identifies issues, illustrates the outcomes of model choice, and identifies the basics of best practice projection methods.

Variable annuities. What is gained – and lost – by recalculating the annuity at regular intervals based on new projections of life expectancy? One way to minimize the projection error is to reestimate and recalculate annuities up to a fixed age of, for example, 80–85. This section identifies the policy issues viewed against the background of the effects on individual and distributional outcomes within the pension cohort.

Socioeconomic determinants of life expectancy in the context of creating the annuity. There is mounting evidence of an increasing gap in life expectancy between persons with lower and higher lifetime earnings. This reflects directly on the outcomes of pensions in all the various forms of earnings-related pension schemes, and not the least defined contribution (DC) schemes – both financial (FDC) and nonfinancial (NDC) ones. Secondary markers of this phenomenon are education and occupation, and in the gender dimension the gap between a low-income and a high-income man can be twice the gap between a low-income and a high-income woman. This section identifies the policy issues arising from these characteristics of the pension pool and discusses the pros and cons of various proposals for dealing with these issues in the construction of the annuity pool(s).

The paper ends with a general summation of the main policy lines developing from the issues discussed.

2. The NDC annuity

The points of departure in formulating an annuity model are (i) the total capital balance on an individual's account at retirement, (ii) an estimate of the life expectancy of the average member of the individual's birth cohort at an age relevant for claiming a pension, and (iii) an assumed future rate of return. This section addresses the issues involved in integrating the rate of return into the construction of the annuity to create ex ante an affordable benefit that also maintains long-term financial sustainability.

2.1. The basic annuity formula

The annuity is a function of the known individual account balance at the individual's chosen age of retirement and projected values of two factors. The first factor is the cohort-based average life expectancy, LE , for cohort k (person j 's birth cohort) at retirement. The second is the internal rate of return, α , computed over the length of average life expectancy.

The account balance is the sum of the individual's contributions up to time τ accredited with a yearly return based on the scheme's internal rate of return prior to the time of retirement. The amount on balance at the end of the period prior to retirement, $K_{\tau-1}$, is divided by an annuity factor, G . The annuity factor includes the average life expectancy at the chosen pension age projected for individual j 's birth cohort and a rate of return α (to be discussed), where $\alpha=1$ if the entire rate of indexation is left outside and applied ex post on a yearly basis. This gives the initial value of the annuity, P_{τ} , for the j^{th} person:

$$(1) \quad P_{j,t} = \frac{K_{j,t-1}}{G[LE_k, \alpha(LE_k)]}$$

Section 3 is devoted to the calculation of life expectancy. This section focuses on the rate of return in the NDC and the pros and cons of different "models" for distributing it over time – either endogenously within the calculation of the annuity or exogenously through its external regular indexation.

2.2. Policy considerations regarding the rate of return to be included in the annuity calculation

In an FDC scheme, the rate of return on the savings transferred to the pool of pensioners is the nominal financial rate earned on the investment of this money in the financial market; i.e., it consists of two components – the real rate of return (r^*) and the rate of inflation (p). Even in a financial scheme it is quite possible to set $\alpha = 1$ and to index the annuity with, for example, a moving average of yearly returns. This is a straightforward way to deal with the uncertainty in the discount rate, as determined by the rate of return.

In an NDC scheme, the rate of return is the internal rate of return, determined by the growth of the nominal contribution wage base (W), which can be broken down into three components: (i) the rate of growth of the real wage per capita (w); (ii) the rate of growth of the number of contributors that underlies this wage growth (λ); and (iii) the rate of change in prices used to deflate the nominal wage to obtain the real wage – i.e., the Consumer Price Index (CPI), or “ p .” In an NDC scheme, the overall rate of return is thus $(1 + w)(1 + \lambda)(1 + p)$, whereas it is $(1 + r)(1 + p)$ for an FDC scheme.

To begin, it is difficult to argue in favor of including the rate of growth of the number contributors (λ) in the computation of the annuity, principally because the determinants are subject to factors that are extremely difficult to project. The rate of growth of the contributing labor force is determined by three factors: (i) the total fertility rate over an average of years two decades earlier; (ii) a country’s current net immigration of prime working-age population; and (iii) individuals’ decisions regarding their formal supply of labor.

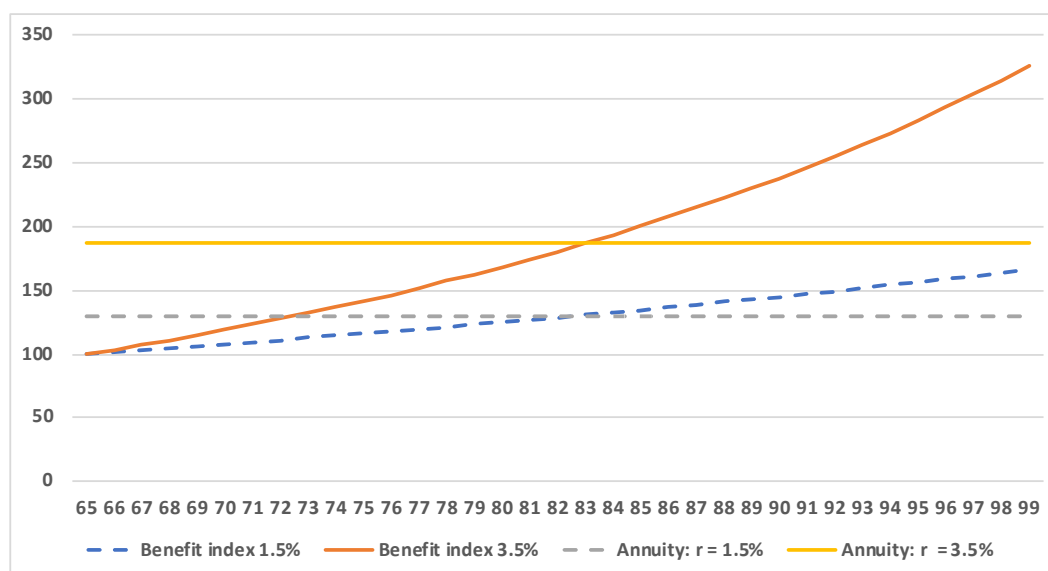
It is highly arguable that positive indexation amounts created by positive growth in the labor force, with λ thus taking on a positive value, should be used to build up a reserve that can be used to finance periods with negative values of λ when the labor force is declining. In periods with a declining labor force – e.g., due to fertility rates under 2.1 or net emigration of the working-age population – the normal cyclical development of the reserves can cushion the need to adjust pensions downward, signaled by temporarily negative values of λ . With a demography with continuously low fertility and net migration from the country (e.g., such as that by characterized Southern, Central, and Eastern Europe since the mid-1990s), positive reserves can only develop through increases in the density of contributions of a declining population. This is clearly possible for emerging economies as they develop economically – the mechanism is explained in Larsson, Leyaro, and Palmer (2019) and for more developed economies in the European Union in Palmer and Stabina (2019).

Alternative values of α can be considered potential candidates for inclusion in the annuity. One alternative is to set $(1 + \alpha) = (1 + w)(1 + p)$ and another is to set $(1 + \alpha) = (1 + w)$, where both of these alternatives exclude the factor $(1 + \lambda)$, as just argued. The latter is the alternative

adopted by the three NDC countries – Italy, Norway, and Sweden. Sweden complements this design feature with an exogenous correction mechanism that accounts for the difference between the norm (the long-term expected value of the return in the formulation of the annuity) and the actual yearly outcome, which then becomes an additional component of exogenous indexation. The argument for leaving out the rate of inflation is also very strong because it is a very difficult parameter to estimate and is more easily employed exogenously (i.e., after it has occurred). Finally, the remaining alternative is to set $(1 + \alpha) = 1$, which means that the rate of return is not included at all in the annuity. In this case, yearly annuity payments are based solely on the individual's capital balance at retirement divided by life expectancy, and indexation of the yearly annuity payment is done exogenously on a regular basis.

For the purpose of illustration, Figure 2.1 assumes an estimate of the rate of per capita real wages $(1 + w)$ of 1.5 percent per year, and with 2 percent inflation, a rate of $(1 + w)(1 + p) = 3.5$ percent.

Figure 2.1: Considerations regarding the rate of return underlying the annuity



Note: Initial benefit of 100 at age 65 with given life expectancy. Inflation = 2%.
Two cases: real return of 1.5% and nominal return of 3.5%.

For comparison, the payment profile of the annuity is also calculated excluding the rate of return, but with each yearly payment indexed (exogenously) with the same assumed yearly rate of return – as it occurs. By including the rate of return in the annuity – with a fixed payment per payout period – the annuity redistributes the time path of payments from the final half of the average life to the first half, and implicitly from those with longer-than-average lives to those with shorter-than-average lives.

2.3. The payment profile over time: A distributional issue

The distributional issue emerging in Figure 2.1 is worth pondering. To begin with, moving a larger percentage of a given sum of money to consumption when younger accords with the belief that people may generally derive greater utility from having a larger part of a given sum of money to consume now rather than later. On the other hand, leaving the indexation outside the creation of the annuity and instead indexing benefits exogenously on a yearly (regular) basis has one obvious advantage: by increasing with the rate of growth in the real wage of all contributors, the relative value of benefits to wages remains essentially unchanged over time. In addition, it results in relatively more income for consumption of goods and services toward the end of life for the very elderly. For those for whom this is the only source of income, this can be important. Figure 2.1 also demonstrates that shifting the lifetime income from a pension to a pensioner's younger years increases the risk of relative poverty in old age.¹ This observation puts emphasis on the importance of an accompanying means-tested minimum income guarantee for pensioners.

2.4. The risk for the insurer

In the payout (annuity) phase of financial old-age pension schemes, because of the uncertainty of the development of the financial rate of return over the 30+ years of life of the members of the pension pool, it is prudent to invest in safe assets with low yields. In addition,

¹ In a study assessing the causes of relative poverty among single elderly women in Sweden, Nelsson, Nieuwenhuis, and Alm (2019); Palmer and Könberg (2019) stress the overall importance of the Swedish means-tested minimum income housing cost guarantee for the elderly in this context.

the annuity provider will keep a large portion of the overall pool of pension balances as secure reserves to cover actual outcomes that diverge unfavorably for the insurer – despite having made a very conservative assumption about the future rate of return.

Obviously, considerable uncertainty exists about the rates of growth of the real per capita wage, the growth of the contributing labor force, and the rate of inflation. In the context of the universal public NDC pension scheme, the obvious way to avoid this risk is – at most – to include only a portion of the expected rate of growth of productivity in the computation of the annuity, where 1.5–2.0 percent is considered relatively high. The rate of growth of the labor force and rate of inflation should also be left outside the annuity computation.

In summary, the reasoning here speaks strongly in favor of either (i) leaving the internal economic rate of return completely outside the creation of the annuity, or (ii) including a conservative estimate of long-term productivity growth – together with the “Swedish” adjustment index (described above) to even out deviations from this “norm.” The adjustment index, the price index, and the labor market index should be employed on a regular (e.g., annual) basis through continual ex post indexation. This is particularly important for an NDC scheme exposed to continuous population deflation (e.g., a chronically low fertility rate) that is not compensated for by positive labor force growth (e.g., through net immigration). Countries with continuous net labor force growth can create a demographic reserve as a buffer to an anticipated cyclical downturn (instead of transferring the dividend directly to current pensioners).

2.5. The role of life expectancy in creating affordable, sustainable NDC pensions

Life expectancy is one of the key parameters of NDC schemes, entering the picture in five important ways:

- Together with individual account balances, it determines the amount of a yearly pension.
- Knowledge of the effect of increasing life expectancy incentivizes individuals to postpone retirement.

- Participants' awareness and behavioral adjustment to changing life expectancy at the micro level increase the amount of the yearly benefit payment over what it would have been if participants exited the labor force earlier.
- The macro result of individuals postponing retirement is an improvement in the labor supply and the economic dependency ratio.
- Life expectancy is used for automatic indexation of the minimum pension.

Finally, life expectancy – together with the average number of years the participant is in the labor force – is a determinant of the time an amount of contributions is in the system before it has to be paid out (the turnover time). *Ceteris paribus*, increasing life expectancy means a longer turnover time – which increases the system's liquidity, reflected as an increase in the present value of assets in the solvency ratio (Palmer 2013). Although to date only Sweden has an automatic balancing mechanism based on a solvency ratio of the nature Sweden has chosen, it is important to keep in mind this technical function of life expectancy projections in the construction of the solvency ratio.

3. Projecting life expectancy

Life expectancy is one of the key parameters of NDC schemes. To begin with, through its role in determining the yearly amount paid out during retirement it contributes to microeconomic efficiency by signaling the need for participants to postpone retirement with increasing longevity. This signal is reinforced by indexing the minimum age to it. In addition, the macro result of individuals postponing retirement is an improvement in the labor supply and national income (and welfare).

Finally, *ceteris paribus*, increasing life expectancy means a longer turnover time from the average age at which a given amount of contributions enters into the NDC scheme and the average age at which it will be paid out. This increase in liquidity increases the "contribution assets" in the system, and thereby has a positive effect on solvency (Palmer 2013). Although to date only Sweden has an automatic balancing mechanism based on a solvency ratio, it is

important to keep in mind this technical function of life expectancy projections in the construction of the solvency ratio.

3.1. Different approaches to projecting life expectancy

By definition, no method for projecting uncertain outcomes can avoid producing errors. In the case of projecting mortality rates, the linchpin of life expectancy modelling, the survival of some cohorts will be overestimated and that of others underestimated. The goal of modelling is that the errors should be random with an expected value of zero over a large number of projections. However, if estimates are systematically over- or underestimated, this leads to a systematic transfer of benefit payments not entirely covered by the sum of money in the pension pool to younger cohorts, over time creating an increasing systematic deficit, financial unsustainability, and an unfair transfer of debt to coming generations.

This section assesses projection methods currently employed with the help of current literature and reports on a new estimation approach that promises better estimates of life expectancy at retirement that, however, suggest their major underestimation by the current dominating approaches – period and cohort models.

3.1.1. The period method for projecting life expectancy

Countries quite commonly do not employ sophisticated statistical projection procedures in making official projections of life expectancy. Instead, there is a long legacy of employing straightforward period statistics based on mortality rates by age in the most recent year – or a moving average of such calculations for a series of years (Bengtsson et al. 2018). This procedure is referred to as the period method in the literature.

The period method gives the average length of remaining life at a given age, assuming that people are going to experience the same age-specific mortality rates observed in the transition from each age x to age $x+1$, beginning with the most recent year one can observe from the actual data. The obvious problem with this method is the “risk” that the decline in mortality rates systematically speeds up over a series of decades (or slows down). The result

is that life expectancy estimates based on the period method are biased toward underestimation.

Using birth cohort data for 1900–2014 obtained from the Berkeley Mortality Database, Alho, Palmer, and Bravo (2013) and Alho, Palmer, and Zhao de Gosson de Varennes (2019) examined the development of cohort mortality from ages 40 and above of 10 countries (Denmark, Finland, France, Italy, Japan, the Netherlands, Norway, Portugal, the United Kingdom, and the United States). Although the profiles of countries vary over the more than a century of birth cohort data examined, what the data show is that (declining) mortality for older age groups is prevalent from the 1960s or shortly thereafter for all the countries examined.

Using a five-year moving average of period data, Alho, Palmer, and Zhao de Gosson de Varennes (2019) show that the period method systematically underestimates the remaining life expectancy at age 65 of the 1,600 completely deceased cohorts that could be followed to age 99 up to 2014. For Sweden, this method yields an error that is as high as 8 percent for the last cohort that turned 65 – which was in 1979 – and that had expired completely in 2014. Applying the period method to data for cohorts turning age 65 from 1980 and forward in time (i.e., younger cohorts that had not yet reached age 99 in 2014) suggests that the gap between projections and outcomes using the period method is increasing as we approach present time. For Sweden, illustrated in a comparison of methods toward the end of this section, the period method underestimates life expectancy from age 65 by on average 8 percent for the cohort that was 65 years old in 1975; i.e., the birth cohort 1910 for all countries examined.

3.1.2. The statistical approach for cohort modeling – the Lee-Carter model and its variants

In 1992, Lee and Carter published what was to become known as the Lee-Carter (LC) model, and a new era of stochastic modelling of mortality and life expectancy arose. The 1992 LC model interprets the log-transformed age-specific mortality at a certain period with two age-dependent parameters and one time-dependent parameter. The model has only one period effect, which means that the underlying assumption is that the rate of change in mortality is time-invariant, as explained in Girosi and King (2007). Although the LC model has proven to

be superior to the period model because of the design feature (restriction) discussed in Girosi and King, it is not flexible enough to fully capture the development of mortality that became evident in the decades beginning in the 1960–1970s, according to the historical data examined by Alho, Palmer, and Bravo (2013).

Many attempts were made to improve the original LC model in the years following its publication in 1992 (Wilmoth 1993; Lee and Miller 2001; Booth et al. 2002; Hyndman and Ullah 2005; de Jong and Tickle 2006; Li and Chan 2007). However, the model extensions of these and other LC variants do not explicitly address the restrictive assumption of time-invariant change in mortality identified by Girosi and King. To the authors' knowledge, the single exception is Booth et al. (2002), who attempted to deal with the time invariance issue by fitting the model with an "optimal estimation period." This has the obvious drawback of having worked for a single (and best) period but not a large number of periods and over a sample of countries.

Booth et al. (2006) compared the model improvements of Lee and Miller (2001), Booth et al. (2002), Hyndman and Ullah (2005), and de Jong and Tickle (2006) using data for 10 countries. They found that the newer variants of the LC model reduced the mean error of log death rates as compared to the original LC-model. However, a word of caution is that lower mean errors of log death rates do not necessarily translate into lower errors in projecting life expectancy over a period of 35+ years – which this paper's work focuses on.

In a study commissioned by the US Office of the Government Actuary, Waldron (2005) analyzed official US long-term mortality projections as well as projection models employed by several European countries. She concluded that regardless of the choice of procedure chosen, ex post evaluations reveal systematic underestimation of expected lifespans. In a series of studies by Alho, Palmer, and Bravo (2013), Zhao de Gosson de Varennes (2016a, 2016b), Palmer and Zhao de Gosson de Varennes (2018), and Alho, Palmer, and Zhao de Gosson de Varennes (2019) analyzed altogether data from 10 countries that cover the period 1900 to 2014. They found that the degree of systematic underestimation escalates with the accelerating improvements that, in varying degrees and more or less pronounced at different

times, have characterized the mortality profiles of cohorts born since the 1950s. Alho, Palmer, and Zhao de Gosson de Varennes (2019) find that for Sweden the underestimation of cohorts born 1935–1944 using the period model is most likely to be more than 10 percent and increasing.

3.1.3. The data analytical period-cohort approach: The Palmer-Alho-de Gosson (PAD) model

Alho, Palmer and Zhao de Gosson de Varennes (2019) present a new projection method for life expectancy, the Palmer-Alho-de Gosson model (PAD model). The PAD model is a “data analytical period-cohort” approach. Its unique feature is that no underlying assumption is made on the rate of change in mortality. Instead, the rate is used as a key parameter for the purpose of the projection as it is what links the period mortalities with the cohort mortalities.

The Appendix contains a formal mathematical description of the model, but an intuitive explanation of the method follows. Assume that the task is to project the remaining life expectancy for those who turned 65 years old in 2015. The data at hand are the age-specific mortality rates for all ages between 65 and 99 from 1900 to 2014. The projection is made by performing the following steps:

Step I: A sequence of empirical rates of change in mortality of past cohorts is attained by using the observed period and cohort mortalities using the second formula in the Appendix, called the PAD model.

In this example, the rates of change in mortality for the cohorts who turned age 65 from 1900 to 2014 are derived using their empirically observed cohort mortality. Note that for the cohorts that turned 65 from 1900 to 1979, the actual rate can be computed as they have completed their lifespan up to age 99 by year 2014. For the later cohorts beginning with those that turn 65 in 1980, only a quasi-empirical rate using incomplete cohort information can be calculated. However, these rates can still be valuable for the purpose of the projection since they provide the most recent information of how mortality has been changing. Alho, Palmer, and Zhao de Gosson de Varennes (2019) present alternative ways of using these “partial”

empirical rates. This section continues with the description of the model of a full cohort (i.e., a cohort with no remaining survivors).

Step II. The rate of change in mortality of the projection cohort is extrapolated based on the series of empirical rates derived in the first step.

A robust extrapolation method should be chosen in this step. The first question is whether to employ a sophisticated statistical method (e.g., an ARIMA model) or a simple extrapolation based on the average rates during the last X number of years. The obvious limitation with ARIMA modelling is that it requires a long time series of around 40 years or more, data that are generally not available in most countries. A second and related question is whether the extrapolation should be based on a long or a short historical time series. Obviously, a short time series from the nearest preceding periods bases estimates on the most recent information. At the same time, it may be wrongly influenced by short- or medium-term and temporary divergence from a longer trend in mortality caused by short-lived events.

Zhao de Gosson de Varennes, Palmer, and Alho (2016) and Alho, Palmer, and Zhao de Gosson de Varennes (2019) test four extrapolation methods, including three simple extrapolations with: (i) the most recent year (PAD-1); (ii) 20 years (PAD-2); (iii) 5 years (PAD-3); and (iv) ARIMA modeling (PAD-4).

Step III: The projection of cohort mortality – and life expectancy – is then calculated with the observed period mortality and the extrapolated mortality rates estimated in the previous step.

In this example, the projection is made using the period mortality (i.e., age-specific mortality in 2014; see the introductory words above to Steps I-III) and the estimated rates for the cohort that turned 65 in 2015.

In the studies of Alho, Palmer, and Zhao de Gosson de Varennes (2019), the data-analytical period-cohort method as illustrated above is examined in considerable detail. The PAD-model with the four different extrapolation methods was compared with the LC model for 1,600 entirely expired cohorts from eight countries: Sweden, Denmark, Norway, France, Italy, the

Netherlands, the United Kingdom, and the United States. The ex post evaluations shown that the longer estimation method (PAD-2 with a 20-year memory) and the ARIMA method (PAD-4) deliver the smallest errors and show practically no tendency toward either systematic over- or underestimation, whereas the two other alternatives are systematically biased toward underestimation, as the discussion above suggests is to be expected.

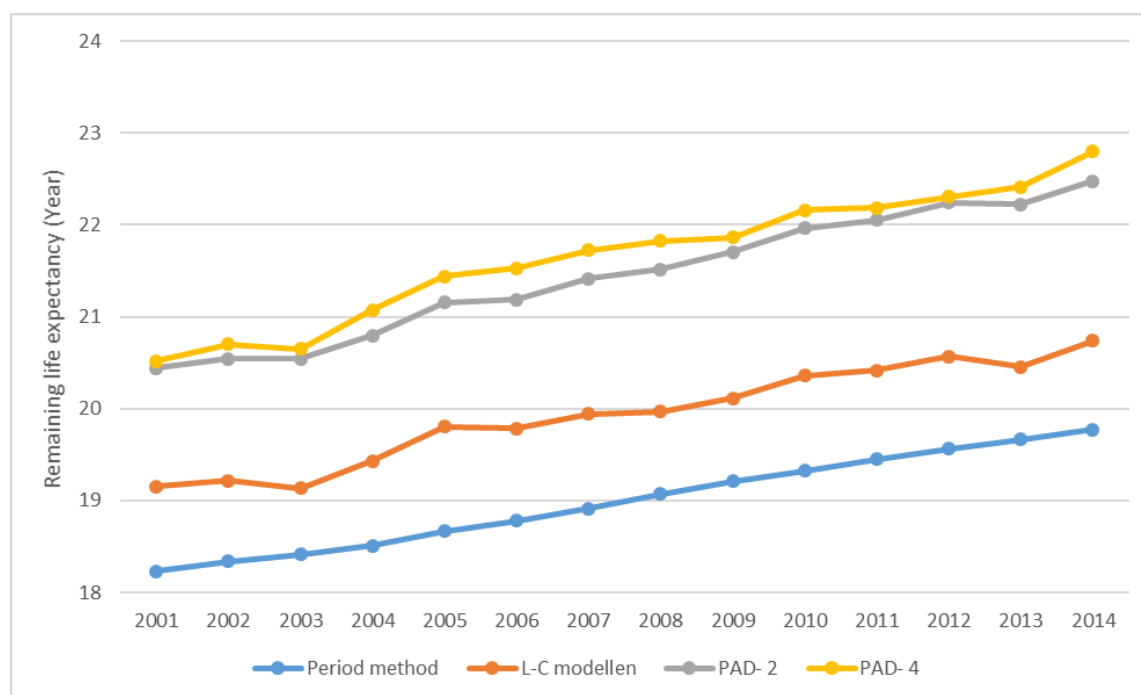
3.2. A comparison of the period model, the LC model, and the PAD model for contemporary birth cohorts

The ex post evaluations examined in the studies by Alho, Palmer, and Zhao de Gosson de Varennes (2019) and summarized above end with the birth cohorts that turned 65 in 1979. The interesting question is, what can be said about the life expectancy of those birth cohorts that have turned age 65 since 1979? To answer this question, another 1,000 individual cohorts from the eight countries could provide a tentative answer – tentative because the youngest cohorts still have a good portion of their lives to live – and in the most interesting cohort, the age group 85–99. To shed some light on this, life expectancy for these cohorts is estimated using the available information and the projections compared with the known outcomes as these cohorts age from 1980 through 2014. These projections constitute what is called the ex ante evaluation in the following.

The results of the analyses in Zhao de Gosson de Varennes, Palmer, and Alho (2016) and Alho, Palmer, and Zhao de Gosson de Varennes (2019) are summarized as follows. First, the period, LC, and PAD (2 and 4) models all estimate a continuous increase in the remaining life expectancy at 65 and older for all birth cohorts from the eight countries. However, the PAD-models almost always give higher projections than the LC model. The exceptions are a few US cohorts.

To relate this result to an NDC public pension scheme, Sweden is chosen as an example. The alternative projection models are applied to persons who turned age 65 in the years 2001 to 2014. The results are shown in Figure 3.1.

Figure 3.1: Projected cohort remaining life expectancy at age 65 for Swedes who turned 65 in 2001–2014



Source: Alho, Palmer, and Zhao de Gosson de Varennes 2019.

The experimental results of this research give a clear warning signal. The acceleration in the rate of decline in mortality is causing not only demographers (e.g., Statistics Sweden) but also practitioners to systematically underestimate the rate of increase in life expectancy. In addition, the strength of the systematic underestimation is increasing with more recent new entrants into the Swedish public pension system. A conservative estimate is that the current Swedish NDC pension scheme is expecting deficits that accumulates to around SEK 95 billion for 10 birth cohorts (those born from 1938 to 1947) (Palmer and Zhao de Gosson de Varennes 2018).

3.3. Policy implications

The conclusions from theoretical literature (Giroi and King 2007) and the empirical literature – and especially the empirical results in the succession of studies by Alho, Palmer, and Bravo (2013) and more recently Alho, Palmer, and Zhao de Gosson de Varennes (2019) – are as follows:

Public pension administrations that are responsible for providing annuities need to move toward more sophisticated projection models backed by evidence that projections are not systematically downward biased under the present conditions of steadily improving mortality rates at and after age 65.

The period model should be abandoned as a projection model for public pension schemes. Given that the assumption of time invariance in the rate of change in mortality is believed to be fulfilled, the LC model works well. However, the empirical evidence shows it is seldom the case that the data fit this assumption. A more prudent approach is to adopt a PAD model that by incorporating more available cohort information into the database used for extrapolation is generally at least as good as the LC model or better. The PAD model improves the likelihood of more accurate and systematically unbiased estimates of the life expectancy of countries' pool(s) of pensioners.

The consequence of choosing a projection method that systematically underestimates life expectancy is the risk of financial unsustainability and intergenerational transfers of debt to the younger generation.

Finally, the research in this area by Alho, Palmer, and Zhao de Gosson de Varennes demonstrates that there is an alternative – the PAD model – that has been tested experimentally and has been proven to work. Moreover, thinking in policy terms, an estimator – such as the period method – that is from the outset known to be unfailingly systematically biased breaks with the underlying principles of NDC schemes.

4. Variable annuities: What is gained and lost by recalculating the annuity at regular intervals based on new projections after the “normal” retirement age?

Current practice in calculating benefits/annuities in public DC pension schemes – including all current NDC schemes – is to employ a once-and-for-all projection of remaining life expectancy at the “normal” pension age.

This practice needs to be reconsidered in the context that mortality in older ages is declining at an accelerating rate in many developed countries and is moving up into the older age groups. This is a main result of Alho, Bravo, and Palmer (2013), which also examines aging in Japan, the leading (larger) nation in aging.

An alternative strategy would be to periodically adjust the annuity (i.e., the stream of yearly benefits) with revised projections of mortality/remaining life expectancy up to a certain ceiling age (e.g., 80 or 85 years old). A variable annuity is similar to adding a new (negative) factor into the indexation of annuities as the pensioner ages.

This sort of approach was suggested by Piggot, Valdez, and Detzel (2005) and Valdez, Piggott, and Wang (2006) for private insurance (but not in conjunction with the PAD model). Alho, Bravo, and Palmer (2013) illustrate its application (without the PAD method of creating the database) in a public NDC pension scheme.

A variable annuity construction means that the risk of deficits caused by underestimation of life expectancy is reduced at the expense of those who survive to the next revision point. This also implies a redistribution within the insurance pool from those who live longer to those who live shorter lives.

Alho, Bravo, and Palmer (2013) demonstrate the expected result that reprojecting remaining life expectancy brings the estimations closer to the actual outcomes, using the LC model as the underlying projection method. Zhao De Gosson de Varennes (2016b) evaluated the variable annuity construction starting at age 65 using the new PAD model, comparing results with a fixed annuity at age 65 based on an LC projection, evaluated with 208 cohorts from Sweden, Denmark, Norway, France, Italy, the Netherlands, the United Kingdom, and the United States. In the variable annuity model life expectancy is reprojected up to age 85 using the PAD database and extrapolation models PAD-2 and PAD-4. Table 4.1 shows that the variable annuity scheme always gives better results compared to the fixed annuity scheme with the same projection model. All reduce the size of the financial deficits.

Table 4.1: Variable and fixed annuities: Average size of financial deficits/surpluses with new projections at 5-year intervals up to age 85 (% of the total capital of a cohort annuity pool)

	Fixed annuity with PAD-2	Fixed annuity with PAD-4	Variable annuity with PAD-2	Variable annuity with PAD-4
Sweden	3.69%	1.98%	2.47%	0.67%
Denmark	3.64%	1.42%	1.97%	0.91%
Norway	1.60%	2.66%	1.28%	0.95%
France	1.84%	4.36%	0.77%	0.50%
Italy	1.78%	5.65%	1.49%	1.69%
Netherlands	3.14%	4.35%	1.87%	0.49%
United Kingdom	1.44%	5.91%	1.24%	1.38%
United States	2.51%	10.33%	3.41%	0.98%

Source: Zhao De Gosson de Varennes 2016b.

Adjusting annuities with new projections of cohort remaining life expectancy to reduce the uncertainty (error) in projections brings a clear improvement in financial stability. The conclusion regarding intra- and intergenerational fairness is that there are both pros and cons involved with the shifting of financing within the pension pool by introducing a variable annuity up to a ceiling age of, for example, 80 or 85. The adjustment of individual younger cohorts' pensions is spread out among many participants, whereas the adjustment is "heavier" per capita for the older cohorts. This causes some loss in the expected utility (consumption) of the oldest, compared with the alternative of an annuity fixed at an earlier date (Zhao de Gosson de Varennes 2016b). At the same time, however, those who gain are a third party (i.e., the younger cohorts), because the transfer of whatever remaining debt there might be is lower. If a reserve fund is established, as was indicated in section 2, then it is likely that this would buffer against fluctuations.

Finally, this observation does not suggest that "the" benchmark age for the pension system should not be indexed to life expectancy. The explicit and transparent presence of life expectancy in the calculation of NDC pensions is a key factor in the automatic stability generated by public NDC pension schemes. And, generally speaking, against the backdrop of

continuously increasing life expectancy, it is important to periodically adjust (through indexation) both the minimum pension age as well as the age at which life expectancy is fixed after some number of revisions.

5. Socioeconomic determinants of life expectancy in the context of creating the annuity

Recall that NDC is a longevity insurance that reflects the lifecycle earnings of participating individuals: “You get what you pay for, with interest.” The DC principle is economically efficient as it rewards formal work in the labor market, which in a universal public pension scheme must be the basis upon which pension rights are attained if the universal scheme is to be viewed as fair. In this perspective, NDCs are both intra- and intergenerationally fair. A logical corollary is that a country’s income distribution policy is preferably pursued through social measures exogenous to the (N)DC scheme design that are created for specific events, such as childbirth, unemployment, sickness, and disability.

What are the ramifications of knowledge of socioeconomic differences in life expectancy within this framework? Section 5.1 presents an overview of the evidence of the correlation between socioeconomic factors and life expectancy. Section 5.2 addresses gender as a separate issue and discusses the reasoning behind the use of unisex life expectancy, in the perspective of gender income inequality. Section 5.3 discusses the pros and cons of some technical approaches proposed in the literature to redistribute pension pool resources, given that this is the chosen policy. Section 5.4 addresses the question of whether gender pooling is a sufficient policy to deal with the gender gap in pensions, which is largely due to the gender gap in earnings.

5.1. Higher life expectancy accompanies higher income, higher education, and choice of occupation

The existence of socioeconomic gaps in life expectancy is by now a well-documented phenomenon. A large number of empirical studies show significant gaps in life expectancy associated with individuals’ position in the income distribution, their level of education, and

their occupation. The general picture is that individuals with higher income, higher education, and professional and white-collar jobs can expect to live longer. This is a feature of life expectancy in countries with relatively high Gini coefficients such as the United States (Gini = 0.4), but also characterizes Western European countries, with generally more equitable income distributions (Gini \cong 2.5–3.0). This section begins with a selection of empirical studies that illustrate the evidence predominantly from the United States, Germany, and Sweden.

Begin with two comprehensive studies of US data. The first is the National Academies of Sciences, Engineering, and Medicine (2015) study of US data, which provides a comprehensive overview of the issues and empirical evidence focusing on the relationships between life expectancy and the distribution of income. The overall conclusion of this study, based on a considerable body of empirical evidence, is that a gap in life expectancy is associated with income (and education) and that it has been increasing over time. The picture is characterized by at best static life expectancy in the lower income deciles, with increasingly longer life expectancy as one moves up the percentiles of the income distribution.

The second study is that of Chetty, Stepner, and Abraham (2016), who analyzed a sample of 1.4 billion person-year observations for individuals aged 40 to 76 residing in the United States. The study focused on explaining the gap(s) in life expectancy associated with income in the United States. The database also includes 4.1 million male deaths and 2.7 million female deaths. One of this study's many interesting results is that the richest 1 percent of women live on average 14.6 years longer than the poorest 1 percent of woman and the richest 1 percent of men live on average 10.1 years longer than the poorest 1 percent of men.

Life expectancy for low-income individuals in local communities was also found to be higher with a higher fraction of immigrants, with a higher fraction of college graduates, and with the level of local government expenditures on community services. For the population as a whole, life expectancy for individuals in the lowest income quartile was significantly correlated with lifestyle risk factors such as smoking, use of narcotics, and obesity; however, Chetty, Stepner, and Abraham (2016) found no significant correlation—either positive or negative—with access

to medical care, physical environmental factors, overall relative income status of the local community, or the local labor market.

No other study matches that of Chetty, Stepner, and Abraham (2016) in breadth of data collected on explanatory factors. However, numerous other country studies support the conclusions of these two US studies. For example, using German data Reil-Held (2000), Klein and Unger (2002), and Lampert, Kroll, and Dundelberg (2007) found that the difference in life expectancy from birth between the lowest and highest income groups is 8.4 years among women and 10.8 years among men. Also using German data, Kroh et al. (2012) found a difference in life expectancy at age 65 between low- and high-income individuals of 3.5 years for women and 5.3 years for men.

Using Swedish data for 1970–2007, Eriksson et al. (2014) found that income inequality in life expectancy, already notable at age 35, increased during the period studied, primarily driven by improvement for high-income persons. These results agree with the other available studies of country data surveyed in other recent studies.² Finally, a more recent Swedish study from Statistics Sweden (2016) found that already at age 35 the difference in life expectancy between the highest and lowest income group is 8 years for men and 4 years for women. Coupling this result to the most significant factor found in Chetty, Stepner, and Abraham (2016) (i.e., lifestyle factors), it is perhaps not a far stretch of the imagination to claim that men are more disposed to take on negative lifestyle habits at early ages – where the prime habits in many countries are still alcohol, substance abuse, smoking, and the absence of regular exercise.

From these and other studies, the chain linking education to occupation to earnings appears to represent different aspects of the same situation – often starting with low or discontinued education. In a study of Swedish white- and blue-collar workers, Zhao de Gosson de Varennes (2016b) found a clear occupational gap between the life expectancy of “nonmanual” and

² Ayuso, Bravo, and Holzmann (2017a) and Zhao de Gosson de Varennes (2016b) also contain recent and extensive overviews of the literature on socioeconomic heterogeneity in life expectancy, supporting the line of thought presented here.

“manual” workers, whereby education is the most important factor in the Swedish context already at age 40 and continuing to 60. Similar results were found by the UK Office for National Statistics (ONS) based on life expectancy at age 65 in England and Wales; however in the context of Finland, the data indicate stagnation in the development (Myrskylä, Leinonen, and Martikainen 2013).

Summing up, gaps in life expectancy are strongly correlated with income. Being in the lower income percentiles is characterized by stagnant life expectancy, while life expectancy is generally higher for persons in higher income percentiles. The empirical evidence suggests this is a “global” phenomenon, at least within the community of highly developed economies for which there are data and analysis of these data.

5.2. Gender differences in income, life expectancy, and NDC annuity construction

A point of departure here is that equal treatment of the genders requires the use of unisex life expectancy in universal public pension schemes, and for that matter, logically, all pension schemes. This is in fact prescribed by the European Union for all of its members. The economic rationale for using unisex life expectancy is straightforward: doing so transfers the accumulated balances of all participants in the pool of pensioners – usually segmented in practice by birth cohort – from the generally richer “class” of men who have shorter lives as pensioners to the generally poorer “class” of women who live longer lives.

Sweden provides an example. A study performed by the Swedish Pensions Agency (2016) divides earnings data for the entire working-age population into five income classes, with each income group also divided by gender. The main result is that 80 percent of the population in low-income groups comprises women. That is, the unisex life expectancy divisor used in the creation of the life annuity at retirement counteracts the regressive profile of the gender income distribution; this in turn is dominated by women in the first four deciles who characterize the profile prior to the redistribution resulting from use of the unisex divisor. The Swedish Pensions Agency’s report also identifies low-income men as the biggest “double losers” – with on average low pensions and low life expectancy.

Before proceeding, consider what lies behind the gap in the yearly earnings of women and men. To begin with, the gender wage gap (i.e., the average difference between the yearly remuneration of working men and women) is 14 percent for Organisation for Economic Co-operation and Development (OECD) countries.³ This does not account for the gender difference in the allocation of time to participation in the formal labor force, attributable to time devoted principally to child care up to adolescence. The Swedish pension gender gap (in 2013) based on the average of all pensioners – just the public pension and excluding the minimum pension guarantee – was 33 percent.⁴ With a gender wage gap somewhat better than the OECD average hovering around 10 percent, this suggests a gap component of around 20 percent, representing less time devoted to work in the formal labor market.

Using NDC individual accounts from 1960, Klerby, Larsson, and Palmer (2019) study the NDC pension accounts of spouses giving birth to at least one child for mothers born 1955–1970, with NDC account data for the period 1960–2012. They find a seemingly fixed pattern for the development of contributions noted on NDC accounts following the birth of the mother’s first child. Not surprisingly, a large earnings gap with respect to the spouse’s yearly earnings occurs from birth and then falls gradually from age 2 of the first child up to age 4 around when the second child is born – giving a period of about six years of coverage by the tax-financed child-care rights (going almost exclusively to mothers). From this point the earnings gap between the mother and her spouse declines, reaching a “steady state” at about 20 percent (on average) until the second child has reached age 12 about 17 years after the birth of the first child. The authors attribute this steady-state earnings gap to parents’ revealed preference for devoting a portion of the couple’s time to care for children at home well into adolescence – behavior well in line with established cultural norms, but it is the mother who supplies the larger part of unpaid care labor at home. On this basis they argue in favor of defaulting account sharing between parents in universal public NDC (and FDC) schemes.

³ Using data available from 2014–2017. <https://data.oecd.org/earnwage/gender-wage-gap.htm>

In summary, computing the annuities in universal public DC schemes using unisex life expectancy makes sense. It results in a redistribution within the pension pool from men with predominantly higher income (and pension balances at retirement) and shorter lives to women with lower income (and pension balances at retirement). Since a substantial portion of the gender income gap is attributable to “structural cultural factors” – occupational gender wage discrimination and women’s part-time work during child-rearing years – the use of unisex life expectancy contributes to evening out the income outcomes – but this is only a second-best policy since it is contingent on the gender difference in life expectancy. With convergence of men’s and women’s life expectancy, this form of redistribution would cease, although the earnings and pension gaps would remain.

5.3. Other approaches to dealing with the socioeconomic gap in life expectancy in creating (N)DC annuities

To reduce the unintended redistribution caused by the socioeconomic gaps in life expectancy, Nalebuff and Zeckerhauser (1985) proposed creating a separate pension plan for each life expectancy group. Esö and Simonovits (2003), Simonovits (2006), Esö, Simonovits, and Toth (2011), and Bommier, Lerous, and Lozachmeur (2011) derived a benefit-age rule under which information available for individuals enables the application of individual-related retirement ages. They showed that this can dampen unintended inefficient redistribution and improve overall social welfare – of course, one might add, given that the rule is fair and representative.

Putting all other considerations aside, an obvious drawback arises in calculating annuities with individualized projections: this requires extensive administration – much of which is sensitive and may not be forthcoming – and presumes that a reliable and manageable statistical procedure exists for “allocating” the individualized life expectancies to create fairly segmented subgroups within the birth cohort pool. In addition, this detailed process may still lead to biased outcomes (among other things because of asymmetric information and adverse selection), and, generally, there is no easy check on the fairness of the procedures chosen until a large number of cohorts have passed through the 35- to 40-year process to the cohort’s extinction.

A more practical approach is to segment the universal NDC pension pool into occupational groups. These have a fairly strong link to both education and income, where income is the most dominant identifier of heterogeneity. This makes sense if a country already has unions with occupational affiliations that are more or less all-encompassing. Some countries already have this sort of segmentation (e.g., Denmark and the Netherlands). Zhao de Gosson de Varennes (2016b) examined theoretically and demonstrated numerically with Swedish data the potential effects on intra- and intergenerational redistribution of a “blue–white” collar segmentation of the population. She found that de-pooling the NDC scheme by occupation can reduce the intragenerational transfer from the blue- to the white-collar occupations by around 5 percent. On the other hand, occupational groups tend to be gender-dominated. For example, occupations with predominantly female workers (of which there are many examples) could no longer benefit from substantial transfers from men within the insurance pool.

Given a decision to segment, Holzmann et al. (2019) showed that a Total Absolute Tax and Subsidy Indicator can be used to compare the extent to which alternative segmentation designs reduce distortions. For example, these authors showed that “de-pooling life expectancy by gender reduces distortions/improves efficiency, but further increases the gap between men’s and women’s pension levels.” The study also demonstrated that a two-tier contribution structure (as in Korea) is able to eliminate most of the heterogeneity effects with a simple and adjustable process as heterogeneity develops. Ayuso, Bravo, and Holzmann (2017b) presented a tax-subsidy mechanism for redistribution within the pension pool and explored alternative policy designs to apply to the new pension cohorts during accumulation. Examples are individualized life expectancy and a two-tier contribution rate structure.

In summary, options and “quasi-options” to using unisex life expectancy were presented herein. The simplest is the quasi-option of segmenting the population into occupational groups at retirement. This implies implicit gender segmentation as well as segmentation by level of education – by separating male-dominated from female-dominated occupations, higher-income from lower-income occupations, and the higher educated from the lower educated. Institutionally, this is the smoothest alternative to implement, but it not only

relinquishes but in fact runs against the goal to reduce the gender gap in pensions through the use of unisex life expectancy.

5.4. To what extent does unisex pooling level the gender playing field?

This section asks the question, “Does unisex pooling level the gender playing field in the end?” The bottom line is that the use of unisex life expectancy embodies an aggregate transfer from men as a group – with higher average income but with shorter lives – to women with longer lives but on average lower income – but it does so in the aggregate. The answer thus is that although it does this on an aggregate average basis as long women’s life expectancy dominates men’s, it will not set right all the “wrongs” on an individual basis.

Given the same average gender gap in earnings prior to constructing the annuity using unisex life expectancy, if life expectancy is exactly the same for both genders, then the lifetime pension gap between genders will remain unaffected. The conclusion is thus that unisex life expectancy succeeds in reducing the pension on average because of the gap in life expectancy.

Taking a simple example, with the population divided equally between men expected to live 18 years on average and women 22 years, the unisex life expectancy factor is 20 years, and provides a 10 percent higher yearly benefit for all women than putting men and women into separate pools. Importantly, this simple example also illustrates the basic difference between a nonfinancial defined benefit (NDB) scheme (e.g., the German or French point system) and an NDC scheme (e.g., the Swedish, Latvian, or Norwegian schemes), which employs unisex life expectancy by virtue of its construction. For the same population, aggregate sum of lifetime contributions, and internal rate of return, NDC with unisex life expectancy contributes toward leveling out the gender income (i.e., pensions) gap, whereas no such built-in mechanism exists in the NDB context.

Taking this one step further, continuing to use Sweden as an example, recent work by Klerby, Larsson, and Palmer (2019) finds that the pension account of a Swedish mother born in 1970 reflects an estimated average earnings gap of 20 percent vis-à-vis the father of her child and

somewhat more for the Swedish standard of almost two children per mother (Klerby, Larsson, and Palmer 2019). The gap originates in time with the birth of the mother's first child. The statistical evidence suggests it can be attributed to the culturally determined preference of parents for one of them (predominantly the mother) to be away from work part-time and at home to care for children up to adolescence (with two children this means a period of about 17 years from the birth of the first child) – where the father is predominantly the full-time worker in the formal labor market.

With a Swedish gender earnings gap of around 20 percent on average and a four-year gap in gender life expectancy, the 10 percent unisex life expectancy factor bonus is a significant step forward in creating gender earnings equality on average for all mothers. But by no means does it deal directly with the cause of the earnings – and consequently pension – gender gap. It will give some individuals too large a “bonus” and others too small a bonus. The drawback of relying on unisex life expectancy to reduce the earnings gap is that it does not deal directly with the ex ante reasons for why women have lower income than men on average. If the life expectancy gap were to disappear, the earnings gap would still remain.

Two pension policy measures can set things straight. The first is (default) sharing of (N)DC pension accounts during a period of shared parenthood. The second is creation (by default) at retirement (of the youngest partner) of a joint annuity at retirement. This is a specific insurance product based on joint income and individual life expectancies, where the surviving spouse (usually the female partner) receives more than a 50 percent share (e.g., 60–65 percent) of the couple's total pension income at the death of the other partner. Both of these social policy measures redistribute pension rights within the domain of the family, instead of indirectly transferring an implicit tax on all men in the pension collective to finance a general transfer to women who have been parents, both those who shared informal care time and those whose division of home care of children was lopsided. Unisex life expectancy is thus a weak second best as a measure for sharing, and it lacks the possible efficiency of sharing of accounts in providing an increased incentive for equal sharing of time between care in the home and work in the formal labor market.

The knowledge that life expectancy is stagnant at the lowest income levels and increases with the level of education and income is another issue. The signal comes at early ages – age 30 and perhaps earlier – as discussed in section 4.1. This illustrates the importance of national and community policy focusing on the first two decades of life and renewed opportunities later in life, together with a focus on acquiring and renewing skills to meet the demands of the labor market, where social support and institutions have an important role to play. This brings to the table the key issues brought out in Chetty, Stepner, and Abraham (2016) regarding lifestyle and longevity, underscoring the need for conscious public health interventions and community services.

Finally, the evidence is overwhelming of the importance of, first, recognizing and, second, dealing through public health efforts the lifestyle issues underlying the low life expectancy of low-income individuals – observable in the data already at ages 30–40. The absence of public health policy initiatives to address these issues is also notable.

6. Conclusions

This paper presented and discussed issues regarding the policy choices that have to be made in determining the construction of the NDC annuity in four areas. All have to do with the criteria of a good universal pension scheme and the backbone of NDCs – affordability, financial sustainability, and intra- and intergenerational fairness. An important message is that the topics of this paper are not all “simply” technical issues; they also have sociopolitical ramifications. The key conclusions from each section follow.

6.1. Construction of the annuity and options for inclusion of the rate of return

This section weighs the advantages and disadvantages of inclusion of the rate of growth of productivity (i.e., partial) in the ex ante calculation of the annuity. The case is made for including a conservative estimate of future long-term productivity growth (i.e., per capita real wages) in the ex ante creation of the annuity – accompanied by an ex post “Swedish” adjustment index based on the difference between actual per capita real wage growth and the “norm” included in the annuity divisor. In addition, the rate of inflation index and the

labor market growth index should be employed on a regular (e.g., annual) basis as a part of the ex post indexation.

With chronic negative labor force growth, this negative growth component must be factored into the overall indexation (usually only reducing the scale of positive indexation), where the ex post model is preferable. In countries with consistently positive labor force growth, a demographic reserve fund can be established to buffer cyclical, hypothesized changes. Finally, the Latvian model (Palmer and Stabina 2019) of dealing with economic dips originating from deeper recessions provides an example of how short bouts of (potential) negative indexation can be smoothed.

6.2. Unbiased estimation of cohort life expectancy

Generally speaking, two criteria for determining the choice of projection method are that: (i) the expected value of projection errors over a succession of birth cohorts is zero; and (ii) the method chosen delivers the highest degree of accuracy as measured by the variation around the mean outcomes. That is, the estimator that comes closest to the actual outcomes (has the lowest random errors) is the most preferable, given that it neither systematically over- nor underestimates ex post outcomes.

With evidence from eight (and in two additional) countries, section 3 shows empirically that mortality rates have generally declined at an accelerating rate through age 85. Judging from the example of Japan, this process can be expected to continue even higher up into the 90s. The challenge for projection methods is to capture this upward movement. This paper shows that the two methods frequently employed by public agencies to project life expectancy at standard retirement ages – the period model and the LC model – systematically underestimate life expectancy at retirement.

A third model, the PAD model, developed by Alho, Palmer, and Zhao de Gosson de Varennes is presented. This model uses the changing relationship between period and cohort mortalities as the basis for projections (the PAD model). It is presented and the three models are compared with the period and LC models using Swedish data. The conclusion from the

literature –illustrated by the Swedish example – is that the period model is not appropriate for projecting life expectancy. Neither is the LC model because it presumes a time-invariant rate of change in life expectancy and as a result is likely to lead to systematically underestimated life expectancy projections. On the other hand, strong evidence is presented that supports the use of the more general PAD model.

6.3. A variable annuity?

This section examines the consequences of not fixing the value of life expectancy to be used in the fixed annuity as early age as 65. Of course, this argument will become partially self-fulfilling as countries index pension ages to life expectancy, where this juncture in the process could follow suit. However, a variable annuity, recalculated at specific ages with reforecasted remaining life expectancy, is similar to adding a new (negative) factor into the indexation of annuities that shifts the cost burden more in the direction of those who live longer. It can be argued that to the extent this leads to a marginal increase in poverty in old age for single pensioners (consisting of an increasing proportion of women as the cohort ages), this can be compensated for through adequate means-tested minimum income supplements – financed with general tax revenues.

6.4. Socioeconomic heterogeneity in life expectancy?

National statistical agencies and researchers are finding an increasing gap in life expectancy with respect to income. It is almost stagnant in the lowest income deciles, but increases progressively up into the higher income deciles. The empirical evidence demonstrates that the gap reflects level of education, occupation, and the resultant income, and the differences are traceable back to as early as age 30, whereas the predominant cause of early death often has to do with lifestyle factors.

The pension outcomes observed reflect the many underlying structural socioeconomic characteristics that contribute to individuals' lifecycle labor market outcomes. It is already well-recognized that the way to address stagnant life expectancy for the least economically well-off is policy aimed at achieving full participation in secondary education, promoting new

skills learning in support of job mobility, and targeting lifestyle issues through better public health programs, also from younger years.

Most importantly, an important gender dimension enters into the discussion of income and life expectancy. This has to do with the fact that women are overrepresented in the lower income quintiles – given structural differences in the density of their labor force participation and the unequal gender division of nonmarket household activities, particularly caring for young children, and labor market gender wage discrimination.⁵

Use of unisex life expectancy goes approximately halfway in narrowing the pension gender gap – from about 20 percent to about 10 percent. And, for example, if the Swedish life expectancy gap of four years were to disappear, the income gender gap in the overall pension collective would still remain.

In conclusion, if the goal is to even out the distribution of shares of the pension pool between men and women the best pension policy measures would be to (i) default sharing of pension accounts between spouses/parents from the birth of the first child, and (ii) default joint annuities, together with unisex life expectancy. This has the desirable effect of evening out the difference in income between spouses and, thus, between genders. These measures work through reducing the earnings gap, and by definition the gender pension gap. These are not the only measures needed, however. The empirical evidence emerging from large-scale studies in the United States shows that the gap in life expectancy has a great deal to do with negative individual lifestyle habits and the availability of important community services in this context. This suggests that policy focusing on the factors underlying the determination of life expectancy outcomes at earlier ages could go far to reduce the causes of low life expectancy among the economically worse off.

⁵ Proponents of unisex life expectancy rest their case on the fact that women's longer lives result in an automatic transfer from male to female participants in the pension pool, which reduces the pension gap.

References

- Alho, Juha, Jorge Bravo and Edward Palmer. 2013. "Life Expectancy and Annuities in NDC." In *NDC Pension Schemes: Progress and Frontiers in a Changing Pension World Volume 2 – Gender, Politics and Financial Stability*, ed. Robert Holzmann, Edward Palmer, and David Robalino, pp. 395–436. World Bank: Washington, DC.
- Alho, Juha, Edward Palmer, and Yuwei Zhao de Gosson de Varennes. 2019. "Projecting Cohort Life Expectancy from the Changing Relationship between Period and Cohort Mortalities." Manuscript in preparation for publication, available from the authors.
- Ayuso, Mercedes, Jorge Bravo, and Robert Holzmann. 2017a. "On the Heterogeneity in Longevity Among Socioeconomic Groups: Scope, Trends, and Implications for Earnings-Related Pension Schemes." *Global Journal of Human Social Sciences-Economics* 17(1): 33–58.
- _____. 2017b. "Addressing Longevity Heterogeneity in Pension Scheme Design." *Journal of Finance and Economics* 6(10): 1–21.
- Bengtsson, Tommy, Nico Keilman, Juha Alho, Kaare Christensen, Edward Palmer, and James W. Vaupel. 2018. "Introduction." In *Old and New Perspectives on Mortality Forecasting*, ed. Tommy Bengtsson and Nico Keilman, Chapter 1, pp. 1–22. Max Planck Institute of Demography: Springer Publishing.
- Bommier, Antoine, Marie-Louise Lerous, and Jean-Marie Lozachmeur. 2011. "Differential Mortality and Social Security." *Canadian Journal of Economics* 44(1): 273–289.
- Chetty, Raj, Michal Stepner, and Sarah Abraham. 2016. "The Association Between Income and Life Expectancy in the United States, 2001–2014." *JAMA* 315(16): 1750–1766.
- Eriksson, Karin, Markus Jäntti, Lena Lindahl, and Jenny Torssander. 2014. "Trends in Life Expectancy by Income and the Role of Specific Causes of Death." Tech. Rep. 8/2014, Swedish Institute for Social Research, Stockholm University.

Esö, Péter, and András Simonovits. 2003. "Designing Optimal Benefit Rules for Flexible Retirement." Technical Report 1353, Northwestern University, Evanston, IL.

Esö, Péter, András Simonovits, and János Toth. 2011. "Designing Benefit Rules for Flexible Retirement: Welfare vs. Redistribution." *Acta Oeconomica* 61(1): 3–32.

Giroi, Federico, and Gary King. 2007. *Understanding the Lee-Carter Mortality Forecasting Method*. Santa Monica, California: The Rand Corporation.

Holzmann, Robert, Jennifer Alonso-García, Heloise Labit-Hardy, and Andrés M. Villegas. 2019. "NDC Schemes and Heterogeneity in Longevity: Proposals for Redesign." Social Protection and Jobs Discussion Paper, No. 1913. Washington, D.C.: World Bank Group.

Klein, Thomas, and Riner Unger. 2002. "Aktive Lebenserwartung in Deutschland und in den USA." *Zeitschrift für Gerontologie und Geiatrie* 35: 528–539.

Klerby, Anna, Bo Larsson, and Edward Palmer. 2019. "Bridging Partner Lifecycle Earnings and Pension Gaps by Sharing NDC Accounts." Social Protection and Jobs Discussion Paper, No. 1919. Washington, D.C.: World Bank Group.

Kroh, Martin, Hannes Neiss, Lars Kroll, and Thomas Lampert. 2012. "Menschen mit Hohen Einkommen Leben Länger." Tech. Rep. 38, DIW, Working Paper, German Institute for Economic Research: Berlin.

Lampert, Thomas, Lars Kroll, and Annlena Dundelberg. 2007. "Soziale Unterschiede in Dermrtalitt und Lebenserwartung." *Aus Politik und Zeitgeschichte* 42: 11–18.

Larsson, Bo, Vincent Leyaro, and Edward Palmer. 2019. "Harnessing a Young Nation's Demographic Dividends Through a Universal NDC Pension Scheme: A Case Study of Tanzania." Social Protection and Jobs Discussion Paper, No. 1922. Washington, D.C.: World Bank Group.

Lee, Ronald D., and Lawrence Carter. 1992. "Evaluating the Performance of the Lee-Carter Method for Forecasting Mortality." *Journal of the American Statistical Association* 87: 659–671.

Lee, Ronald D., and Timothy Miller. 2001. "Evaluating the Performance of the Lee-Carter Method for Forecasting Mortality." *Demography* 38(4): 537–549.

Myrskylä, Mikko, Taina Leionon, and Pekka Martikainen. 2012. "Lifespan Variation by Occupational Class: Compression or Stagnation Over Time?" Tech. Rep. 010/2012, Max Planck Institute for Demographic Research: Rostock.

Nalebuff, Barry, and Richard J. Zeckerhauser. 1985. "Pensions and the Retirement Decision." Technical Report, 1285, NBER, Cambridge, MA.

National Academies of Sciences, Engineering, and Medicine. 2015. *The Growing Gap in Life Expectancy by Income: Implications for Federal Programs and Policy Responses*. Committee on the Long-Run Macroeconomic Effects of the Aging U.S. Population-Phase II. Committee on Population, Division of Behavioral and Social Sciences and Education. Board on Mathematical Sciences and Their Applications, Division on Engineering and Physical Sciences. Washington, DC: The National Academies Press.

Nelsson, Kenneth, Rense Nieuwenhuis, and Susanne Alm. 2019. "Sweden: Adjoining the Minimum Pension with NDC." Social Protection and Jobs Discussion Paper, No. 1909. Washington, D.C.: World Bank Group.

Palmer, Edward. 2013. "Generic NDC: Equilibrium, Valuation and Risk Sharing With and Without NDC Bonds." In *Nonfinancial Defined Contribution Pension Schemes in a Changing World. Volume 2: Gender, Politics and Financial Stability*, ed. Robert Holzmann, David Robalino, and Edward Palmer, pp. 309–333. Washington, DC: World Bank.

Palmer, Edward, and Bo Könberg. 2019. "The Swedish NDC Scheme: Success on Track with Room for Reflection." Social Protection and Jobs Discussion Paper, No. 1901. Washington, D.C.: World Bank Group.

Palmer, Edward, and Sandra Stabina. 2019. "The Latvian NDC Scheme: Success Under a Decreasing Labor Force." Social Protection and Jobs Discussion Paper, No. 1902. Washington, D.C.: World Bank Group.

Palmer, Edward, and Yuwei Zhao de Gosson de Varennes. 2018. *Pensionsystemens Demografiska Utmaningar*. SNS Förlag.

Piggot, John, Emiliano A. Valdez, and Bettina Detzel. 2005. "The Simple Analytics of Pooled Annuity Funds." *Journal of Risk and Insurance* 72(3): 497–520.

Reil-Held, Anette. 2000. "Einkommen und Sterblichkeit in Deutschland: Leben Reiche Inger?" Discussion Paper 504, Department of Economics and Statistics, Mannheim University.

Simonovits, András. 2006. "Optimal Design of Pension Rule with Flexible Retirement: The Two-Type Case." *Journal of Economics* 89: 197–222.

Statistics Sweden. 2016. "Life Expectancy and Mortality of Different Socioeconomic Groups." *Demographic Report* 2016:2.

Swedish Pensions Agency. 2016. "Longevity and Group Income." (*Livslängdsanalys för olika inkomstgrupper*). Pensionsmyndigheten: Stockholm.

Valdez, Emiliano A., John Piggott, and Liang Wang. 2006. "Demand and Adverse Selection in a Pooled Annuity Fund." *Mathematics and Economics* 39(2): 251–266.

Waldron, Hilary. 2005. "Literature Review of Long-Term Mortality Projection." *Social Security Bulletin* 66(1): 16–30.

Zhao de Gosson de Varennes, Yuwei, Edward Palmer, and Juha Alho. 2016. "Projecting Cohort Life Expectancy Based on its Rate of Change in Mortality." In *Benefit Design, Retirement Decisions and Welfare Within and Across Generations in Defined Contribution Pension Schemes*, ed. Yuwei Zhao de Gosson de Varennes, Economic Studies 157, pp. 13–48. Uppsala: Economic Studies, Uppsala University. ISBN 978-91-85519-64-4.

Zhao de Gosson de Varennes, Yuwei. 2016a. "Benefit Design, Financial Risk and Inter-Generational Transfer in NDC Pension Schemes." In *Benefit Design, Retirement Decisions and Welfare Within and Across Generations in Defined Contribution Pension Schemes*, ed. Yuwei

Zhao de Gosson de Varennes, *Economic Studies* 157, pp. 49–72. Uppsala: Economic Studies, Uppsala University. ISBN 978-91-85519-64-4.

Zhao de Gosson de Varennes. 2016b. “Socioeconomic Differences in Life Expectancy, Unintended Intra and Inter-Generational Transfers in DC Pension Schemes.” In *Benefit Design, Retirement Decisions and Welfare Within and Across Generations in Defined Contribution Pension Schemes*, ed. Yuwei Zhao de Gosson de Varennes, *Economic Studies* 157, pp. 111–148. Uppsala: Economic Studies, Uppsala University. ISBN 978-91-85519-64-4.

Appendix. A formal description of the PAD model

To formalize the model, the notation is first defined.

Age-specific mortality rate at age x and year t is given by:

$$m_{x,t} = \frac{D_{x,t}}{N_{x,t}}$$

where $D_{x,t}$ is the number of deaths at age x and year t , and $N_{x,t}$ is the corresponding population at risk.

The starting point of the analysis is a calendar year, t , and only ages from some x up to the highest age w are considered. The set of relevant period mortality rates for year t is:

$$M_{x,t} = \{m_{x+z,t} : z = 0, 1, \dots, w - x\}$$

The set of cohort mortality rates starting at $t + 1$ is:

$$M_{x,t+1}^c = \{m_{x+z,t+1+z} : z = 0, 1, \dots, w - x\}$$

The average rate of change in mortality of the cohort can be expressed by the relationship between the period and cohort mortality as follows:

$$\xi_{x,t} = \frac{1}{w - x + 1} \sum_{z=0}^{w-x} \frac{\ln(m_{x+z,t}) - \ln(m_{x+z,t+1+z})}{z + 1}$$

Suppose then that the year $t = T$ is the last year for which period mortality data are available, or known, $M_{x,T}$. The cohort mortality rate, $M_{x,T}^c$, can be predicted in terms of the cohort rate of change in mortality and the period mortality rates:

$$\hat{m}_{x+z,T+1+z} = m_{x+z,T} \exp(-(z + 1)\hat{\xi}_{x,T})$$

where $\hat{\xi}_{x,T}$ is the predicted value of $\xi_{x,T}$. This gives, as a byproduct, an estimate of the period mortality, $M_{x,T+1}$, so the procedure can be repeated using those estimates and the predicted

value $\hat{\xi}_{x,T+1}$ to obtain a forecast for mortality in $M_{x,T}^c$. In this manner, forecasts for $M_{x,T+k}^c$, $k = 0, 1, \dots$ can be obtained iteratively.

Once the predicted values of $M_{x,T+k}^c$, $k = 0, 1, \dots$ are available, remaining life expectancies in age $x + z$, $z = 0, 1, \dots, w - x$ can be computed with the usual methods.

Incompletely observed cohorts at jump-off time

What remains is to formulate the specific model for incorporating information from the predicted life expectancy of birth cohorts that immediately preceded the cohort under study. For example, to estimate life expectancy during the interval $x = 65$ and $w = 99$ at year t , the last cohort with full empirical age-specific mortality is 65 years old at year $t - 35$. For those cohorts that are age 65 between year $t - 34$ to $t - 1$, only partial empirical information of age-specific mortality is available. However, this additional cohort information can still be used. The feasibility of applying linear weights to the observed rates of change in mortality of these adjoining but still incomplete cohorts is examined (i.e., $\tilde{\xi}_{x,t}$, and the projected rates of decline $\hat{\xi}_{x,t}$ for all $t \in [t - 34, t - 1]$). In other words, the weight given to the observed rates of decline of the first incomplete cohort at $t - 34$ is $34/35$, and the weight decreases to $1/35$ for the latest incomplete cohort. The weight given to the projected rate of decline of the first incomplete cohort is then $1/35$ and increases gradually to $34/35$ for the latest incomplete cohort. Formally, the average rate of change of incomplete cohorts can be written as:

$$\xi_{x,T-z} = \frac{z}{w-x} \tilde{\xi}_{x,T-z} + \frac{w-x-z}{w-x} \hat{\xi}_{x,T-z}, \text{ for all } z \in [1, w-x]$$