Netspar

Network for Studies on Pensions, Aging and Retirement

Netspar DISCUSSION PAPERS

Marike Knoef, Rob Alessie and Adriaan Kalwij

Changes in the Income Distribution of the Dutch Elderly between 1989-2020

A Dynamic Microsimulation

DP 09/2009-030 (revised version March 2012)

Changes in the income distribution of the Dutch elderly between 1989-2020: a dynamic microsimulation *

March 2012

*The authors thank Stichting Instituut GAK and Netspar for their financial support. Furthermore, we wish to thank Wim Bos (CBS), Arjan Soede (SCP), and Centrum voor Beleidsstatistiek for tracking down the changes in the IPO data due to the revision in 2000. Finally, we thank Hans-Martin von Gaudecker, Didier Fouarge, Koen Caminada, participants of EALE 2009, participants of the conference 'Labor Force Participation and the Well-being of the 50+ population' in Utrecht (2009), participants of the European Workshop on Dynamic microsimulation modeling in Brussel (2010), participants of the Econometric Society World Conference 2010, and two anonymous referees for their useful comments.

Abstract

This paper analyzes the income distribution of the Dutch elderly using a microsimulation model. Microsimulation models allow for detailed estimates on the whole income distribution. Our model deviates from traditional models by explicitly considering the persistency and heteroskedasticity of real income shocks. In this way, modeling all underlying processes influencing household income becomes less necessary, which can improve the trade-off between refinement and the tractability of microsimulation models.

We show the results of three model specifications with different levels of refinement. The results are in line and indicate that between 2008 and 2020, the highest predicted annual growth among the elderly is for median-income households (about 1.2%). High-income households have a somewhat lower predicted growth (about 1.0%) and low-income households only have a predicted annual income growth of about 0.5%. Inequality therefore seems to increase in the lower part of the distribution, while it will probably decline in the upper part of the distribution.

Keywords: income distribution, population aging, microsimulation.

JEL codes: D3, J11, J14.

1 Introduction

In most developed countries the aging of the population places an increasing financial burden on society through pay-as-you-go financed social security, pension, health, and long-term care systems (OECD, 2011). Since the 1990s, social security programs and pension schemes in many developed countries are, therefore, being redesigned (e.g. Gruber and Wise, 2004).

Policies aimed at alleviating the costs related to the aging society can be based on the notion that the financial burden is shared between generations (see Van Ewijk et al., 2006; Bovenberg and Ter Rele, 2000). Alternatively or at the same time, one could call upon intragenerational solidarity, such as solidarity within the elderly generations.[1] In order to assess the viability of proposed reforms to redesign pension schemes in developed countries, policymakers require insights into the income distribution of current and future generations of pensioners in a situation of no policy changes. It is important to note that also without pension reforms, the future income distribution of pensioners will differ from the current distribution due to developments in longevity and in demographic and socio-economic compositions. For instance, in many countries the number of divorces is increasing and female labor force participation has increased strongly during the last decades, so that many more women will receive occupational pension income in the future. Also, there are productivity differences between cohorts that lead to income differences.

Using a microsimulation model detailed estimates on the future income distribution are possible.[2] Internationally, there are several microsimulation models built for income predictions and pension issues. E.g. the MIDAS (Microsimulation for the Development of Adequacy and Sustainability) model simulates the adequacy of pensions in Belgium, Italy and Germany (Dekkers et al., 2008). Pensim2 for the UK estimates the future distribution of pensioners income and aims to analyze the distributional effects of proposed changes to pension policy (Emmerson et al., 2004). In Sweden the SESIM model, started in 1997, investigated the Swedish national system of study allowances. Since the year 2000 the focus has shifted from education to pensions and the model now studies the income of the Swedish babyboomers and the financial sustainability of the Swedish pension system (Flood et al., 2006). Other examples of microsimulation models that have been constructed mainly because of the growing concern about population aging are DYNASIM3 and the MINT (Modeling retirement Income in the Near Term) model

[1] In the Netherlands an increasing part of the pay-as-you-go public pension scheme is financed by general tax revenues. Consequently, also the 65+ population pays for the state pensions and due to the progressive Dutch tax system, this policy option redistributes income within the elderly generation.

[2] Merz (1991), O'Donoghue (2001) and Zaidi and Rake (2001) explain, review, and classify microsimulation models around the world.

for the US (Panis and Lillard, 1999; Butricia et al., 2001; Toder et al., 1999; Smith et al., 2007), and DYNACAN for Canada (Harding, 2007).

This paper constructs a dynamic microsimulation model to predict the evolution of the income distribution of the Dutch elderly until 2020, taking into account demographic and socio-economic changes. In contrast to previous microsimulation studies, for the income predictions we estimate a fixed effects income equation and we study the income process by explicitly paying attention to the modeling of the error terms. Households may experience income shocks, the distribution of which may be different for different types of households. In addition, income shocks may have persistent effects, and the degree of persistency may vary over the lifecycle. Therefore, we allow for autocorrelation, with the autocorrelation pattern being a function of age.

The advantage of using fixed effects and modeling the autocorrelation pattern of the error terms of the income equation is that it makes it less necessary to explicitly model the underlying processes that determine household income. Yet, more complex simulation models give more underlying information. For example, only after explicitly modeling labor market status, we can say more about the income positions of elderly with and without occupational pension income. Modeling the income process, however, can improve the trade-off between refinement and tractability of microsimulation models.

The dynamic aging approach as implemented in this paper is also applicable to other countries, when analyzing distributional income effects of demographic and socio-economic changes. For illustrative reasons in this paper we use three levels of refinement for the income equation. The first specification only contains age and period effects. It models no other underlying processes that influence income (for example labor market states) and thus relies heavily on the modeling of the income process. In the second specification household demographics are added, and the third specification also incorporates the labor market status of household members. In these specifications changes in demographic variables or labor market status lead to income shocks. The main results of the three specifications are rather similar. From this finding we cautiously conclude that adding other background characteristics will not affect the simulation results dramatically.

The results show that next generations of Dutch pensioners probably have higher equivalized household incomes than current generations of pensioners, especially for median income households. Between 2008 and 2020, equivalized household income of the elderly in the age group 65-90 is predicted to grow on average by about 0.5% per year for the 10th percentile, about 1.2% for the median, and about 1.0% for the 90th percentile. These may not be specifically Dutch trends. For example, also in other OECD countries the female labor force participation has increased strongly in the last decades. Therefore, depending on the pension rules in different countries, women have built up (more) pension rights, which may also in other OECD countries lead to future pension incomes that grow relatively the most for median income households.

If one aims to quantify the effects of different pension policies, then it is important to model labor supply responses explicitly (Creedy and Duncan, 2002). This is beyond the scope of this paper, however. This paper offers insights into the development of the future income distribution, induced by increased longevity and ongoing demographic and socio-economic changes. On the other hand, if labor market outcomes of a certain policy measure are known, they can be incorporated into the model.

This paper is structured as follows: the next section describes the relevant features of the Dutch pension system. Sections 3 describes the data and section 4 presents some descriptive analyses on the income distribution and labor force participation. Section 5 describes the microsimulation model, after which section 6 summarizes the estimation results. Section 7 presents the simulation results and the paper concludes with section 8.

2 The Dutch pension system

As in many European countries, the Dutch pension system consists of three pillars. The first pillar is a pay-as-you-go system and involves a flat rate public pension benefit for all residents as from the statutory retirement age of 65 onwards. Everyone who has lived or worked in the Netherlands between the age of 15 and 65 receives a public pension (also those who do not work). The level of the public pension is linked to the minimum wage and depends on the number of years residing in the Netherlands. Couples who have lived in the Netherlands between the age of 15 and 65 each receive 50% of the minimum wage, and single pensioners receive 70% of the minimum wage. People that have not lived in the Netherlands as from the age of 15 do not receive the full amount of first pillar pension benefits. If they have a very low or no occupational pension and almost no wealth, the first pillar is topped up with social assistance to guarantee a social minimum.

The Dutch second pillar consists of capital funded occupational pensions, of which the primary responsibility lies with employers and employees. Occupational pensions in the Netherlands have a mandatory nature, such that 90% of the employees have a pension scheme with their employer. All people who have earnings above the minimum wage rate build up pension rights, also part-time workers, proportional to their part-time factor. The unemployed and disabled build up some pension rights in most collective agreements. After the age of 65 unemployment and disability benefits from the government stop and people receive their public and occupational pensions just as everyone else. Labor contracts are mostly terminated at the statutory retirement age of 65 according to collective agreements.

Occupational pensions mainly consist of defined benefit pension plans, that are transferable to a surviving spouse. The benefits are determined by pre-retirement earnings, years of employment, and by the rules of the public and private pension systems. Until the nineties most pension plans aimed to pay a pension income of 70% of final gross wage as from the age of 65 if an employee has worked full-time for at least 40 years. From the early nineties onwards pension funds have lowered their ambition, they now aim to pay 70% of the average career salary, instead of 70% of the final gross salary.

The third pillar is formed by private individual pension products. Everyone can buy third pillar pension products to save for extra pension benefits, but they are mainly used by the self-employed and employees in sectors without a collective pension scheme.

3 Data

The data are taken from the 1989-2007 Income Panel Study of the Netherlands (IPO, Inkomens Panel Onderzoek, CBS 2009) and the 1995-2007 population register (GBA, Gemeentelijke Basisadministratie, CBS 2010), both gathered by Statistics Netherlands.

3.1 Income Panel Study (IPO)

The IPO, a representative sample of Dutch households, consists of an administrative panel dataset with income information. Most of these data are from the Dutch National Tax Administration. In the IPO, so called 'key persons' are randomly drawn from the Dutch population and are followed over time. Data on all household members of the key persons are also available. Major advantages of having administrative data are a very low attrition rate and a high level of representativeness. It is a well-known fact that the rich and the poor are often underrepresented in surveys, institutional households are in general not included, and the elderly population and single person households have relatively low participation rates in surveys (Alessie et al., 1990; Knoef and De Vos, 2009). Another advantage of administrative data is that the observed variables are measured with a high degree of accuracy. A drawback of the IPO is that it lacks some crucial background variables, such as education levels. Variables that are included in the data are individual characteristics (such as gender, date of birth and marital status), household characteristics (such as family composition) and financial variables related to income. As from the year 2000 the IPO dataset has been revised because of major tax reforms. Details about this revision can be found in Knoef et al. (2009).

The raw sample consists of 1,835,819 key persons. We remove 1.5% of the sample because of a missing age or a missing or non-positive household income. Furthermore, we exclude households with nine or more household members and households where the key person is a member of a multiple couple household, a child or a student. This selected sample consists of 1,290,226 observations. Then, we select all households where the key person is born between 1917-1970 and is of age 36-90[3], which leaves us with 911,079 observations. Finally, households in the bottom or top 0.1% of the income distribution are regarded as outliers and excluded from the analysis. Because of the revision mentioned above, the year 2000 is presented two times in the data. We keep the year 2000 before revision instead of the year 2000 after revision, since the tax reforms that caused the revision started in 2001. The resulting sample consists of 861,336 observations.

3.2 Population register (GBA)

GBA is the population register in the Netherlands. This register provides, among other things, information on marital status of all people registered in Dutch municipalities.[4]

Data is available from January 1 1995 to January 1 2008. Just as in IPO, we select all persons born between 1917 and 1970 in the age group from 36 to 90 year. Furthermore, as we want to estimate transitions between t and $t+1$, the marital status in $t+1$ has to be known. Therefore, 2006 is the last year we can use and persons who, for example, emigrate or decease in $t+1$ are excluded at time t.

We end up with 6,812,340 individuals in 1995, increasing to 8,673,138 individuals in 2006. Table 1 presents the share of people making a transition in marital status. The percentage of married people who divorce between t and $t+1$ raised from 0.7% in 1995 to 0.8% in 2006. Furthermore, per year on average 2.5% of the divorced persons make a transition into marriage. Most widows and widowers are relatively old and do not remarry again. On average, 0.4% of the widows and widowers make a transition into marriage from one year to the other.

[Table 1]

[3]In this way we can make predictions until 2020 for the population of age 50-90, because the cohort born in 1970 reaches the age of 50 in 2020 and the cohort born in 1917 is of age 90 in the last wave of the data (2007). In this study we ignore new immigrant families. For the elderly we expect the effect of this ignorance to be small.

[4]Individuals not registered as residents are, for instance, NATO personnel, diplomats and individuals illegally residing in The Netherlands.

4 Developments in income and the labor force participation

4.1 Descriptives equivalized household income

Before making predictions about the future income distribution of the Dutch elderly, this section describes developments in the past. We use equivalized net household income, which is defined as the sum of all incomes received by the household, net of taxes and social insurance contributions, measured in 2005 euros using the consumer price index and, for multiple-person households, divided by the equivalence scale provided by Statistics Netherlands (Siermann et al., 2004).[5] Henceforth, any reference to 'income' should be read as 'net equivalized household income'.

The choice of the equivalence scale can affect inequality rankings (Buhmann et al., 2005). We use the equivalence scale proposed by Statistics Netherlands because it is based on the Dutch situation. Kalmijn and Alessie (2008) found that the modified OECD scale and the equivalence scale of Statistics Netherlands yield very similar results. We analyze the distribution of net equivalized household income for all key persons of the households in the sample, as this is a representative randomly drawn sample of the Dutch population (see section 3.1).

Table 2 describes the distribution of income for key persons in the age groups 50-64 and 65-90, respectively. In the age group 50-64, mean income increased by 21%, from 20,114 euro in 1989 to 24,351 euro in 2007. In the age group 65-90, income was fairly constant during the 1990s. It increased by only 1% between 1990-1999, compared to 9% between 2000-2007. This is probably related to the fact that no indexation of public pension benefits occurred in the early 1990s.

[Table 2]

The Gini coefficient and the decile ratios show that inequality in the age group 50-64 increased between 1989 and 1995 and remained fairly constant thereafter. This is in accordance to the results of Gottschalk and Smeeding (2000), who also found a growing inequality in a number of other OECD countries from the mid 1980s to the mid 1990s. Caminada and Goudswaard (2001) found that the two main forces behind this phenomenon are a more unequal distribution of market incomes and changes in social transfers. Furthermore, in 1990 a revision of the tax system led to more inequality. According to SCP (2003) also the growth in the number of two-earner couples increased inequality between 1985 en 1994.

For the age group 65-90, inequality is lower and shows a different pattern. It grew between

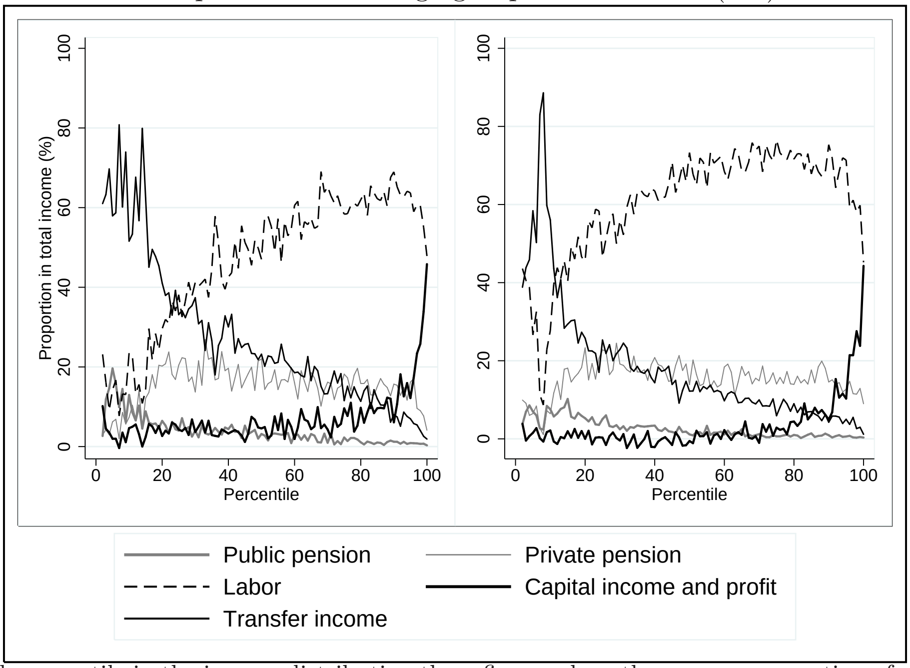
[5] Our income concept takes into account labor income, transfer income, capital income, income taxes, taxes on wealth, social insurance contributions, tax deductible mortgage interest, and the imputed rent (a percentage of the value of an owner-occupied house over which one has to pay taxes).

1989-1991, but declined in the years after 1991. Since 1998, inequality in the age group 65-90 has been quite stable. Several factors may have induced these trends, such as the increased number of women receiving occupational pension incomes, changes in early retirement schemes, the development of the pension system, and the business cycle.

4.2 Income composition

Figure 1 presents the composition of income across the income distribution for the age group 50-64 in 1989 and 2007. The horizontal axes give the percentiles of the income distribution, the

Figure 1: Income composition for the age group 50-64 in 1989 (left) and 2007 (right)

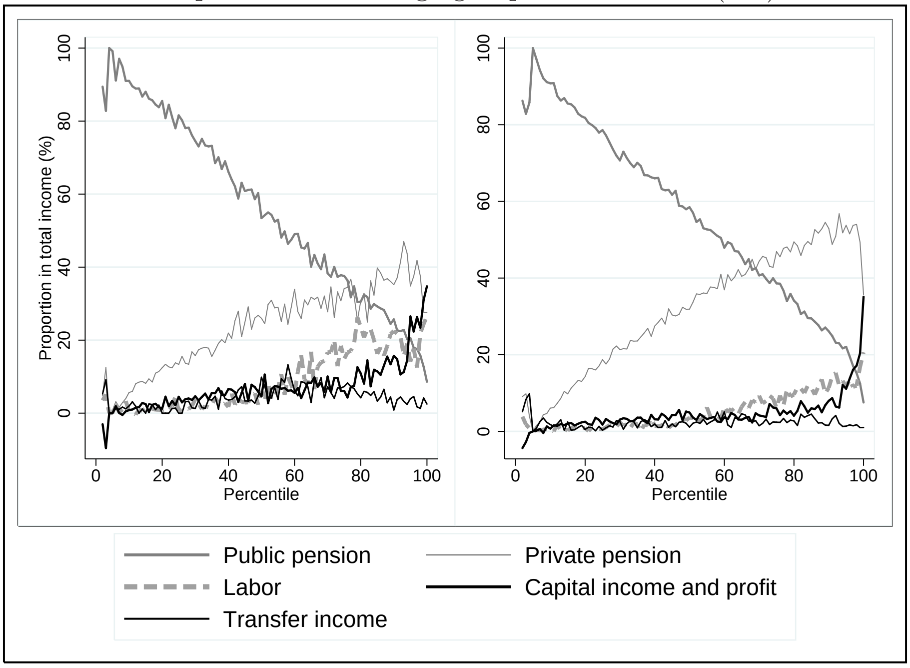


Note: For each percentile in the income distribution these figures show the average proportion of several income components. Transfer income includes welfare, disability benefits, and unemployment benefits. Occupational pensions also contain early retirement income.

vertical axes the proportions of the various income components. For example, for the age group 50-64 we see that the importance of labor income increases over the income distribution. The importance of labor income has also increased between 1989 and 2007 over the whole income distribution, and at the cost of transfer incomes. This is explained by higher participation rates in this age group for both men and women, which will be described in section 4.3 below.

Figure 2 presents the income composition for the age group 65-90 in 1989 and 2007 and shows that the income of a median income household consists of 58% public pension benefits, 32% private pension income, and 10% remaining income sources. The income share of private pensions increased between 1989 and 2007 at almost all percentiles. For example, the income share of private pensions is higher than the income share of public pensions for incomes above

Figure 2: Income composition for the age group 65-90 in 1989 (left) and 2007 (right)



Note: For each percentile in the income distribution these figures show the average proportion of several income components. Transfer income includes welfare, disability benefits, and unemployment benefits.

the 85th percentile in 1989 and for incomes above the 70th percentile in 2007. This finding might be explained in part by spending cuts of the government during the early 1990s, as a result of which public pension benefits have not been adjusted for inflation. In principle, public pensions follow the minimum wage, which is linked to the development of the contractual wages. If no indexation takes place, public pensions will lag behind the growing prosperity; all the more because contractual wages in turn lag behind earned incomes, because of occasional increments and promotions (De Kam and Nijpels, 1995). Other explanations for the increase of the income share of private pensions are the increased female labor force participation and the development of the pension system in the 1950s and the 1960s (Deelen, 1995).

Above the age of 65, labor income is still a significant share of income for households in the high percentiles (although the relative importance diminished somewhat between 1989 and 2007). It is likely that households with the 'better' jobs remain active on the labor market and obtain a relatively high share of their income from work. Capital income only plays a substantial role for the top 5% of the households.

4.3 Labor force participation

Labor is an important source of income and in general leads to pension entitlements for old age. Figure 3 shows the percentage of males and females receiving labor income across several generations. For example, '1938' refers to persons born in 1938. The vertical differences between lines measure the 'cohort-time' effects. We use this terminology to emphasize that it is not possible to disentangle age from cohort and time effects in this figure. Between the ages of 50 and 65 there is a steep decrease in the proportion of males receiving labor income. However, as from the generation born in 1943, more men in the age group 50-65 continue to participate in the labor market. At age 60, about 46% of the men born in 1938 received labor income and about 61% of the men born in 1943 received labor income. This means there is a cohort-period effect of 15%-points, probably due to policy reforms that have been implemented since the late 1990s to discourage early retirement. The results are in line with Kapteyn et al. (2010), who found that the labor force participation for men of age 55-64 decreased until 1993 and has increased afterwards. Furthermore, they agree with SCP (2006), who report that between 1994 and 2003 the importance of labor earnings for the age group 55-64 increased, at the expense of income from early retirement.

Figure 3: Percentage of males (left) and females (right) receiving labor income, by age and cohort

The right figure shows that, as in a lot of OECD countries (OECD, 2004), female participation rates have strongly increased across cohorts and time. This has also been found by Euwals et al. (2011), who claim that changed attitudes towards the combination of paid work and children have played a major role in the Netherlands. This trend will have considerable consequences for the income structure of next generations of pensioners, as more two-earner couples today will lead to more couples receiving double pension incomes in the future. In the data we see that the proportion of two-earner couples increased between 1989 and 2007, from 23% to 40%. On the other hand, the proportion of couples where only the man receives labor income declined from 36% to 16%. The percentage of couples where both men and women receive a private pension income has increased from 1.5% to 6.5%.

With regard to inequality, more two-earner couples can lead to a pooling effect: the inequality within the group of households with two earners is lower than that of couple households with one earner. This means that an increase in the proportion of two-earner households will, at a certain point, reduce household income inequality.

5 Microsimulation model

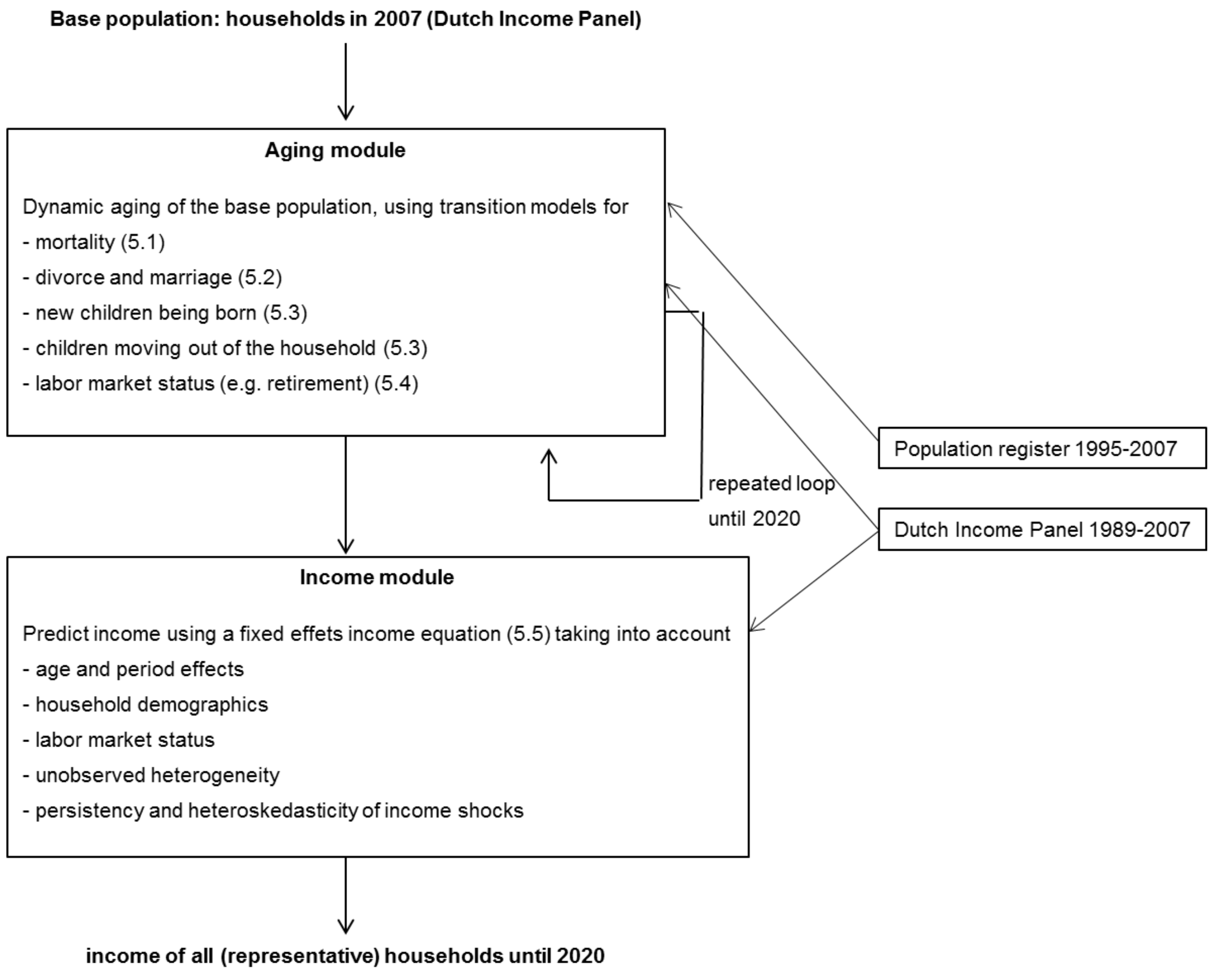
Microsimulation models are used for income predictions and pension issues internationally. These models are in general very demanding multi-year projects with huge data demands (Harding, 2007). One often needs to combine various data sources with different samples, so that one needs to rely on matching of 'statistical twins' (e.g. Geyer and Steiner, 2010) and on surveys that often suffer from representability problems, especially when focusing on the elderly population. For instance, in surveys the elderly population living in private households are often underrepresented (Knoef and De Vos, 2009) and nursing homes are often excluded.

In a microsimulation model the quality of the input data is of prime importance: if the baseline data are not representative, the predictions of the population will not be representative either (Martini and Trivellato, 1997). Our study uses a long and representative administrative panel. Although administrative data contain less detailed information on the characteristics of persons and households, the panel aspect of the data allows us to take into account unmeasured variables such as education, ability, and cohort effects.

To simulate the income distribution of the elderly until the year 2020, we use an open dynamic population model with cross-sectional aging. In the model each characteristic for each person is updated each year (dynamic aging). By contrast, in microsimulation models with static aging individual characteristics are constant over time. Then, the weights attached to each individual change over time and mimic the process of demographic aging. Static aging is well suited for short to medium term forecasts (3-5 year), where it can be expected that large changes have not occurred in the underlying population (O'Donoghue, 2001). An example of a model with static aging can be found in Soede et al. (2004), who analyze future incomes in six European countries. Cross-sectional aging means that we first simulate all individuals for one year, then for the second year, and so forth. Longitudinal simulation models, on the other hand, simulate individual one for all years, the same for individual two, and so forth. Cross-sectional aging allows us to have interactions between household members. For example, husbands and wives make joint labor supply decisions, and the death of a household member can influence the labor market states of the remaining household members. Our model is open, as marriage and birth lead to new synthetic household members. In closed microsimulation models the matching of spouses is restricted to persons within the sample.

Figure 4 describes the design of the model. The representative households in the Dutch Income Panel of the year 2007 form the base population of the model and are the starting point of the simulation. We dynamically age all members of these households until 2020 in the aging module, where people age, they may decease, divorces may take place, children may leave their parental home, new partners or children may enter the household and labor market states may change. We take into account that mortality risks vary between different parts of the income distribution and explain the implementation of this in section 5.1. To predict the transitions in household demographics and labor market states transition models are used, which are estimated with IPO and GBA data. Section 5.2 and 5.3 explain the transition models with regard to marital status and children, while 5.4 describes the transition models with regard to labor market status.

Figure 4: Design of the microsimulation model



After the aging module households move into the income module, where household incomes are predicted using the simulated characteristics of the households until 2020. To this end, we estimate a fixed effects income equation, taking into account age and period effects, household demographics and labor market status. The fixed effects take into account unobserved heterogeneity and we consider the persistency and heteroskedasticity of income shocks. Section 5.5 explains the income equation in detail.

5.1 Differential mortality

In the aging module, where we age all household members in the microsimulation model from 2007 to 2020, persons may decease. To determine whether an individual in the sample deceases or not we apply Monte carlo simulations (see e.g. Law and Kelton, 1982). Therefore, for each individual and at each period from 2008 to 2020 we draw a random value from the uniform distribution. If this random value is lower than the predicted mortality rate the individual deceases. We use predicted mortality rates per age, cohort, and gender published by Statistics Netherlands and adjust the mortality rates of the first and fourth income quartile using the degree of differential mortality found by Kalwij et al. (2011) for the Netherlands. If we would not take into account differential mortality we would underestimate the income level of the elderly, as low income households would survive relatively too often and high income households would survive not often enough.[6]

With regard to the degree of differential mortality Kalwij et al. (2011) find, in line with findings in other European countries (e.g. Von Gaudecker and Scholz (2007) for Germany and Osler et al. (2002) for Denmark) a quartile ratio Q1/Q4 of 2.2 for men and 1.7 for women as from the age of 65. This means that mortality rates in the first income quartile are 2.2 times higher for men and 1.7 times higher for women, relative to the fourth quartile. As from the age of 65 we therefore adjust mortality rates such that mortality rates in the first quartile are 2.2 (or 1.7 for women) times higher than in the fourth quartile, keeping the average mortality rate equal. Before the age of 65 mortality rates are small, such that differential mortality will not make a relevant difference.[7]

[6] When we do not take into account differential mortality, the average yearly income growth of pensioners between 2008 and 2020 is 0.15%-points lower.

[7] To determine which households belong to the first and the fourth income quartile, we use the income position corrected for the age profile (using the fixed effects income equation where only age and period effects are taken into account).

5.2 Transitions in marital status

Using the population register, we model the following transitions in marital status from year to year: married-divorced, unmarried-married, widow(er)-married, and divorced-married. Logit models are employed for men and women separately to estimate transition probabilities between the various marital states. We use age and year of birth as explanatory variables and assume period effects to be negligible compared to the age and cohort effects. We do not explicitly model transitions into widowhood. Becoming a widow(er) depends on the death of a partner. This probability is incorporated via mortality (described in section 5.1).

We assume people to make at most one transition in marital status per year and apply Monte Carlo simulation to assess whether a change indeed occurs. In case of a divorce, the partner of the key person is removed from the household, and in case of marriage a new household member is added. These new household members have the same age as their partners and the opposite gender.

5.3 Transitions in the number of children

The probability of a child leaving the parental home from one year to the other is estimated using a logit model, where age and gender of the child are the explanatory variables. For this estimation we select all children in IPO in 2006 and check whether they are still in the household in 2007. Thus, for the years 2008-2020 we assume children to have the same behavior with regard to leaving their parental home as the children between 2006 and 2007.[8]

The probability of a newborn child in the household is also modeled with a logit model. The explanatory variables are the age and gender of the key person in the household, whether there is a couple in the household, and the number of children which are already present in the household. For this estimation we select all households in the years 1989-2006 and we determine, given the characteristics in $t-1$, whether a new child has entered the household during the next year. The simulation model ignores children already born to enter the household.

5.4 Transitions in labor market status

The model distinguishes three labor market states: (1) receiving labor income, (2) receiving occupational pension income and (3) receiving none of these two ('other'). In order to belong to (1) or (2), labor income or occupational pension income has to be at least 500 euro per year. In case an individual receives both labor income and pension income the highest income

[8] Children are defined as all persons younger than 30 who are at least 18 years younger than the key person of a household.

component counts.[9]

We model the transitions between the three labor market states and assume 'occupational pension' to be an absorbing state. Concerning singles, we estimate multinomial logit models for men and women separately. The labor market states of the two members of a couple are interrelated. For couples we therefore treat the three labor market states of a husband and a wife as $3 \times 3 = 9$ univariate outcomes. For instance, we estimate the transition probability from the state where both husband and wife work to the other eight states with a multinomial logit model. The explanatory variables used in the estimations are age, cohort, marital status, and the number of children.

Using the parameters of the transition models, we estimate the transition probabilities for all singles and couples, given their age, marital status, and labor market status in the previous period. Also here, we use monte carlo simulation to determine whether a transition takes place.

To determine the labor market status at time $t + 1$ with the transition models, we need the labor market status at time t. A problem arises for new household members and children who enter adulthood. To determine an initial state for them we estimate a multinomial logit model per gender, with age and cohort as explanatory variables. The increased labor market participation of women therefore enters the model in two ways: via the initial labor market states of women and via cohort effects in the labor market transition models.

5.5 Income equation

To predict income trends for future generations of pensioners we model household income using a fixed effects model with age, period, and socio-economic variables. Socio-economic variables enable us to take into account developments in the income distribution due to different socio-economic characteristics of future pensioners. The fixed effects allow us to control for time-invariant omitted variables that influence the income of a household. They capture education, ability, and cohort effects that incorporate productivity differences between generations (Kapteyn et al., 2005). Fixed effects are in line with Haveman et al. (2007), who found that pre-retirement economic advantages continue into retirement. We could find one other microsimulation model using fixed effects. That is the MINT model that uses fixed effects to take into account unmeasured heterogeneity in lifetime pre-retirement earning profiles (Toder et al., 1999; Butricia et al., 2001).

[9] The number of people receiving both labor and occupational pension income is low. In the period under consideration a lot of people got incentives to retire (very) early. Gradual retirement was an uneconomic choice and almost did not occur. E.g. workers retiring later than the earliest possible early retirement date were not compensated by higher benefits or lower taxes, so that in fact they faced an implicit tax rate of more than 100% (Kapteyn and De Vos, 1999).

The fixed effect reflects the financial well-being of a household, corrected for household size and other explanatory variables. When the household composition in the simulation changes, most likely because of a divorce or the death of a partner, we assume that, apart from the effect of the divorce or widowhood, the financial well-being of the remaining partner remains the same (we do not change the fixed effect). In reality this is often the case because of widow pensions or alimentation. Also, when people divorce, ex-partners in the Netherlands often receive half of their ex-partners occupational pension after the age of 65. In the model we assume that the percentage change of income due to divorce or widowhood are the same for everyone. Only, when there are systematic differences between income groups regarding mobility, this may affect the simulation results.

The disadvantage of a fixed effects estimator in microsimulation models is that it rules out out-of-sample simulations (Wolf, 2001). However, in this analysis we can use a fixed effects model because our target population are future pensioners, who are already born and available in the data. In the simulation the income profiles are estimated with the same data as the base population is derived from. The fixed effects income equation is

$$y_{it} = \alpha + \beta' x_{it} + \mu_i + v_{it}, \tag{1}$$

where y_{it} is the 'log' of equivalized household income[10] of household i in time period t, α is a scalar, x_{it} is the it-th observation on K explanatory variables, β is a parameter vector of size K, μ_i is the unobserved individual effect and v_{it} is the error term. We have to assume strict exogeneity

$$E(v_{it}|\mu_i, x_{i1}, ..., x_{it}, ..., x_{iT}) = 0 \tag{2}$$

and identify α using the normalization $\sum_{i=1}^{N} \mu_i = 0$. The strict exogeneity assumption implies that there is no feedback effect from income to the explanatory variables. This means, for example, that we have to assume that within a cohort income has no effect on retirement decisions. On the other hand, the influence of the income growth among cohorts on retirement decisions is taken into account by modeling cohort effects in the transition models. The estimation of α, β and μ_i is explained in appendix A.

We estimate three specifications of the income equation with different levels of refinement of the model. In the first specification, the vector x_{it} only contains age and period effects. With this specification income mobility only results from income shocks. By adding additional variables to the vector x_{it}, more individual heterogeneity is introduced in the income path. In the second specification, we add demographic variables such as household size and marital

[10] If one wants to simulate income components separately, one should take into account the correlations between components. There is no need for that in this paper.

status. Including household size as an explanatory variable, in addition to the equivalence scale already used in the dependent variable, leads to information about the income effect of an additional man, woman or child in the household. For example, if the coefficient for the number of adult men in the household is positive, we can conclude that on average, the income of one additional man exceeds his marginal costs of living (determined by the equivalence scale). The third specification also takes into account the labor market states of household members. Whether people are active on the labor market, inactive or retired, influences household income.

Occupational pensions of elderly depend on their labor market histories and we do not, or do not completely, observe these. For example, when young people are unemployed this influences their old-age income. Fortunately, the youth unemployment rate in the Netherlands is low and long-term unemployment (unemployed for more than 1 year) does not often occur.[11] Also, part of the people with unemployment benefits do built up some second pillar pension rights paid by their old employer.[12] Furthermore, the fixed effects control for this as persons who are disadvantaged in the labor market have a relatively low fixed effect which translates in relatively low income during working life and retirement. Nevertheless, the above mentioned data limitations prevent taking into account individual-specific impacts of economic downturns in our microsimulation model.

Age and period effects are implemented as dummy variables, so that their relationship with income is very flexible. However, these age and period effects cannot be identified empirically together with the fixed effects. To identify age and period effects when fixed effects are controlled for, we follow an identification strategy similar to the one of Deaton and Paxson (1994) and assume that all time dummy coefficients add up to zero and are orthogonal to a linear time trend. We thus assume that all period effects are due to unanticipated business cycle shocks.

Households experience income shocks, the size of which may depend on characteristics of the household (heteroskedasticity). For example, income shocks may be larger during working life than during retirement. Furthermore, the question arises how long income shocks persist (autocorrelation), and whether the persistency of a shock depends on age.

When a household experiences an income shock in period t, this may have an effect on the income in the periods following t. The error term v_{it} therefore might follow an autoregressive scheme. To model this we fit the following auxiliary regression model of order two[13]

$$v_{it} = \rho_{1,it} v_{i,t-1} + \rho_2 v_{i,t-2} + \epsilon_{it}, \tag{3}$$

[11] Between 1996 and 2004 the long term youth unemployment rate was on average only 1.47%.

[12] This depends on collective agreements. In the public sector, people who become unemployed still built up 37.5% of their pension rights during the period in which they receive unemployment benefits.

[13] We find that higher orders are of no importance.

where we assume ϵ_{it} to be serially uncorrelated. The persistency of a shock may depend on age. For example, one would expect that income is more smooth over time after retirement, implying a higher autocorrelation of the residual component after retirement than before. Kalmijn and Alessie (2008) provide support for this and report that the two-year autocorrelation of equivalized income is quite stable during midlife, but moves to a higher level after the age of 65. Therefore, we allow $\rho_{1,it}$ to be a function of age.

As explained above, the variance of an income shock may depend on the characteristics of a household. We take this heteroskedasticity into account by investigating the distribution of ϵ_{it} for several mutually exclusive groups of households. For example, the group of households where the key person is younger than 65 and the group of households where the key person is older than 65, for singles and couples. For each group we draw income shocks from the empirical distribution of residuals in 2001-2007 for that group $(\widehat{\epsilon}_{i,2001}, ..., \widehat{\epsilon}_{i,2007})$.[14]

In the predictions we assume period effects to be zero, such that the predicted incomes are free from the effects of the business cycle. Finally, we take into account that as from 2015, a partner bonus for younger partners of state pension beneficiaries with no or low income will be abolished. We subtract the partner bonus for all households who are not eligible for a partner bonus anymore, and of whom the younger member of the couple has no labor income. SZW (2009) found that remaining household income for most of these households will not reach the eligibility limit for social assistance.

6 Estimation results

6.1 Income equation

Table 3 presents the estimation results of the fixed effects income equation for three different specifications. The first specification includes only age and period variables (first and second columns). In the second specification, household demographics are added (third and fourth columns) and in the third specification also labor market states are added (fifth and sixth columns).

[Table 3]

In all three specifications, age effects increase until about the age of 55 and decrease afterwards. As from the age of 70 they increase again, because of selective attrition through mortality. Although we use a fixed-effects model, there is selection related to the idiosyncratic

[14] By using the years as from 2001, possible effects caused by the revision of the data are excluded.

errors. Selective attrition leads to inconsistent estimates, therefore, we cannot interpret the age coefficients as 'causal' effects of age on income. However, for the predictions this does not matter, since this is the best linear predictor of income for those who survive. To let the 'right' people survive we implemented differential mortality in the aging process of the microsimulation model (section 5.1).

The shape of the age profiles for the second and third specification are very similar, while the age profile of the first specification is more pronounced. The first specification has relatively high age effects around the age of 54, caused by children leaving their parental home. Equivalized household income increases when children leave their parental home, as the equivalence scale captures the fact that children cost money.[15] Specification two and three correct for the presence of children, hence they have lower age effects around age 54. The estimated period effects follow the development of the business cycle.

The second specification shows that households with more adults have on average a higher equivalized household income. On average adults thus yield more income than 'costs' (in terms of the increase in the equivalence scale). Households with more children, on the other hand, have on average a lower equivalized income. Kalmijn and Alessie (2008) found that this is mainly due to an increase in expenditures by having children and only to lesser extent due to a decline in the personal income of women after the birth of children.

Marital status is significantly associated with income. Compared to divorced men, divorced women are relatively worse off. A divorce often coincides with a loss of an adult in the household, such that the total effect of a divorce for men is a 2.8% loss of income (0.033-0.061) and for women a 25% loss of income (-0.123-0.131). Widowers and widows are better off than unmarried men and women, and the unmarried are on average better off than divorced men and women. Men have on average 8% more income in widowhood than in marriage, but women are 9% financially better off in marriage than in widowhood.

The third specification takes labor market states into account, namely, the number of men and women receiving labor income, and the number of men and women receiving occupational pension income. Due to the possible endogeneity of labor market status, the estimated coefficients are likely to be biased and therefore we do not allow for a causal interpretation. That is because those people that we observe to be at work are probably the people that have unobserved (time-varying) characteristics that make it relatively profitable for them to work.[16] Nevertheless, the coefficients can be used in a least squares projection. The least squares pro-

[15]Money transfers between parents and children not living in the same household cannot be taken into account because they are not available in the data.

[16]The fixed effects do take into account 'ability', but we have no information about time specific 'good unobservables' that make it relatively profitable (or unprofitable) for people to select themselves into work.

jection is the best predictor in that it minimizes the mean squared error. Hayashi (2000) devotes a section (2.9) on least squares projection. He explains that if the assumptions justifying the large-sample properties of the OLS estimator are not satisfied OLS provides a consistent estimator of the best way to combine linearly the explanatory variables to predict the dependent variable as long as the researcher has a random sample available. As expected, the higher the number of working men and women, the higher is household income. An additional worker in the household increases equivalized household income on average by only 12%, which is related to the high amount of part-time work (especially among women) and the tax system. More household members with an occupational pension also increases household income. According to the model, the net effect of retirement is a reduction of household income by about 2-6%. These high net replacement rates correspond with net replacement rates reported by the OECD (2011).

The parameters $\rho_{0,1}$ to ρ_2 in table 3 show the autocorrelation pattern over the lifecycle. ρ_1 (the first order coefficient, see equation (3)) is a function of age. We experimented with several specifications and also investigated whether it is relevant to specify ρ_2 as a function of age. We found that income shocks are persistent and that persistency increases with age. Only, around the statutory retirement age of 65 income persistency is low, probably because the income composition changes as from that age. In the first specification ρ_1 increases from 0.22 at age 36 to 0.76 at age 90. In the second and third specification ρ_1 is somewhat smaller, especially before the age of 65. This can be explained by the fact that the added demographic and labor market status variables capture part of the persistency. Consider for instance a person faced with a negative income shock from a transition to unemployment. Specification three takes labor market status into account, so as long as the person stays unemployed the negative income effect persists. In the first two specifications labor market states are not taken into account explicitly. However, a person can receive a negative income shock, which may implicitly be caused by unemployment. The parameters $\rho_{0,1}$ to ρ_2 determine the persistency of the shock. This persistency increases with age, comparable to the duration of unemployment, which also tends to increase with age. Finally, σ_μ and σ_ϵ show that the individual variation is larger than the random component.

Future income shocks are drawn from the empirical distribution of the idiosyncratic residuals in the years 2001 to 2007. As shown by Kalmijn and Alessie (2008), the variance of equivalized income (logged) is relatively low after 65. We therefore distinguish between households with key persons younger and older than 65. The standard deviation of the residuals is 40% higher for households where the key person is younger than 65 than for households where the key person is older than 65. In the third specification we also distinguish households that do receive labor or occupational pension income from those that do not receive any of these income components. For households where the key person is younger than 65, the standard deviation of the residual is 49% higher in households without labor or occupational pension income, compared to households with labor or occupational pension income. In households where the key person is older than 65, the standard deviation of the residual is 71% higher for households without occupational pension income, compared to households with an occupational pension. In the simulation these results lead to higher income shocks for young households and for households without labor and/or occupational pension income.

6.2 Transition models

6.2.1 Marital status

Table 4 provides the estimation results of the transition models of marital status. The results show that old persons and persons in old cohorts divorce less often than young persons and persons in young cohorts. Men remarry more often than women after a divorce or the death of a spouse, but with age both men and women less often remarry. After a divorce, young cohorts remarry less often than old cohorts, while on the other hand young cohorts remarry relatively more often after the death of a spouse. Furthermore, with age less people are going to marry, but persons in young cohorts marry more often than persons in old cohorts. This may seem counterintuitive, as it is commonly known that persons in young cohorts marry less often than persons in old cohorts. However, this can be explained by young cohorts marrying later in life than old cohorts. Therefore, in the age group under consideration (36-90) the number of marriages is relatively high for young cohorts.

[Table 4]

6.2.2 Children

Table 5 shows that, as expected, with age more children leave their parental home. Furthermore, daughters leave their parental home earlier than sons. As from the age of 36, when age increases less children are being born. In addition, more children are born in young cohorts (who in general give birth to children later in life) than in old cohorts. Children are more often born in a couple household and in households where already one child is present than in single adult households and households without children. On the other hand, in households with two children or more there are relatively few births.

[Table 5]

6.2.3 Labor market status

Transitions in labor market status are estimated for singles and couples separately.[17] Table 6 presents estimation results for single males and females.[18] The results show that persons in young cohorts keep working longer than persons in old cohorts. The estimated probability for an employed 36 year old single female to stay employed until the age of 60 is 10% for the cohort born in 1940, 20% for the cohort born in 1950, 32% for the cohort born in 1960, and 44% for the cohort born in 1970. Furthermore, divorced men and women experience transitions from work to 'other' and from 'other' to work relatively often, and for women the number of children is positively associated with transitions from work to 'other' (e.g. out of the labor force).

[Table 6]

Finally, table 7 shows the gender-specific estimation results of the multinomial logit model for the initial labor market status of new household members and children who enter adulthood. All household members in the data are used. Labor force participation ('labor') increases until about the age of 40 and decreases afterwards and with age more people receive an occupational pension. Persons in young cohorts more often have a labor or occupational pensions status than persons in older generations.

[Table 7]

[17]The multinomial logit model assumes conditional stochastic independence of the error components of the alternative choices (IIA). We used two commonly used tests, the Hausman test and the Small-Hsiao test, to test the IIA assumption. The test results were inconclusive which means we have no unequivocal information about whether the IIA assumption was violated by our data, the final income results however, appear not to be very sensitive to the inclusion of demographic and labor market transitions (the results of the first, second and third specification are rather similar). Therefore, we cautiously conclude that transition models that relax the IIA assumption, such as the random effects multinomial logit model, would probably not influence the simulated income results very much.

[18]To save space, the detailed estimation results of the labor market transitions of couples are available on request.

7 Simulation results

Corresponding to the three specifications of the income equation, we have three predictions of the income distribution until 2020. Before explaining the income predictions we describe the predictions of marital status and labor market status from the aging module, as they are input for the income predictions in the income module. Predictions of marital status are given in table 8 for the age groups 50-64 and 65-90. In the age group 50-64, the most important finding is the growth in the share of unmarried and divorced people. In the age group 65-90 we find that widowhood among women decreases. This can be explained by the converging life expectancies of men and women, which leads to younger cohorts of women being less often widowed. Furthermore, the fall in widowhood can be attributed to the babyboom generation reaching the age of 65. Therefore, the total age group 65-90 starts to contain relatively many 'young' elderly who are widowed less often (composition effect). The results in table 8 agree with the long term projections of Statistics Netherlands (De Jong and Van Huis, 2003).

[Table 8]

Table 9 presents predictions of labor market status. For both men and women, and both age groups 50-64 and 65-90, the share of people receiving occupational pension income increases. This especially holds for women, as a result of the strong increase in their labor force participation.

[Table 9]

Using the predictions of marital status and labor market status described above, we predict equivalized household income for all households. Table 10 shows the results of the most extensive prediction, where household demographics and labor market states are taken into account (model specification three). When interpreting the results one should take into account that the statistical uncertainty surrounding these predictions may be substantial and we have to make careful statements.[19]

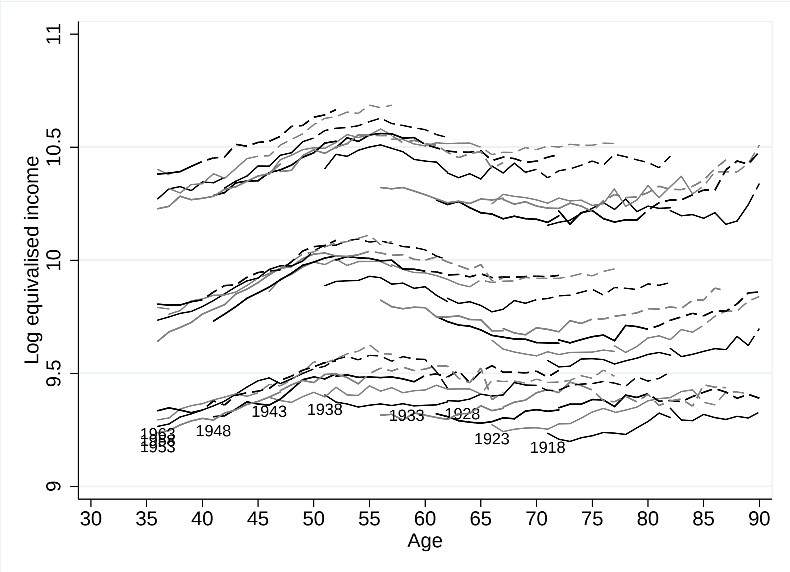
[Table 10]

Incomes in these tables are free from period effects, such as the effects of the business

[19] When we want to compute confidence bands, we need drawings from the parameter distributions from every transition model and for the income equation and evaluate the model again for each draw. This, however, is not feasible, because of the enormous computation time required.

cycle. According to the predictions, income will increase on average by about 0.6% per year for the age group 50-64 and 1.0% per year for the age group 65-90 between 2008 and 2020. The Gini coefficient and the decile ratio p90/p10 show that inequality in the age group 65-90 will probably increase until about 2012 and stabilize thereafter. Focussing on the decile ratios p90/p50 and p50/p10, two contradictory developments seem to occur: an increasing inequality in the lower part of the income distribution and a decreasing inequality in the upper part of the income distribution. This shows the importance of investigating the entire income distribution by microsimulation, rather than just investigating the development of an inequality measure such as the Gini coefficient. Inequality indices differ in their sensitivities to income differences in different parts of the distribution, but one index cannot show the different developments occurring throughout the entire income distribution. For the age group 50-64 the Gini coefficient and the decile ratio p90/p10 show that inequality is predicted to decrease until 2012 but to increase thereafter. After 2012 inequality is predicted to rise in the upper part of the distribution as well as in the lower part of the distribution.

Figure 5: Log equivalized household income per age and cohort.

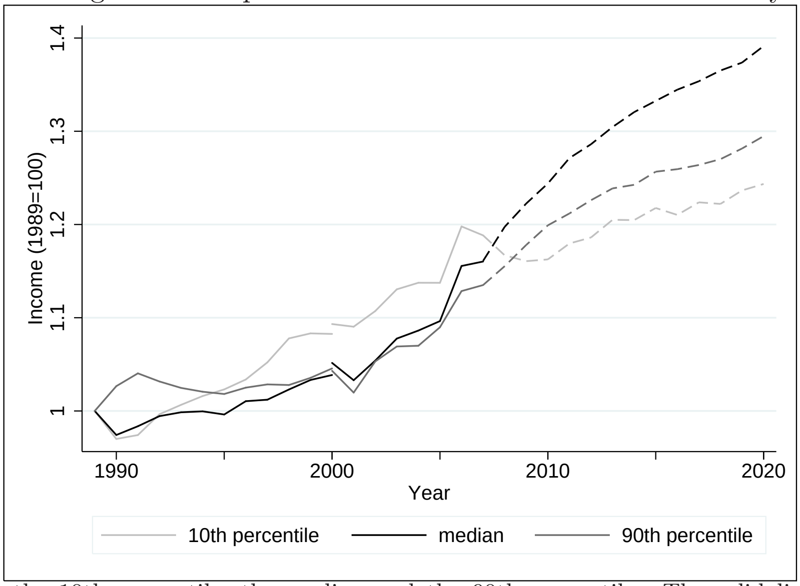


The 10th, median, and 90th percentile of log equivalized household income per age and cohort. The solid lines are realizations corrected for period effects, the dashed lines are predictions made with the most extended version of the microsimulation model (specification three).

Figure 5 shows realizations and predictions of log income per age and cohort. For every cohort the figure presents the income of the 10th, the median, and the 90th percentile. Period effects are excluded for the predictions (the dashed lines) as well as for the realizations (the solid lines). We use *log* equivalized household income, as it is more interesting to compare relative than absolute changes. The age profile of the median incomes and the 90th percentile is stronger than that of the 10th percentile. As expected, young cohorts have higher incomes than old cohorts. However, for the 10th percentile cohort-time effects decrease between 2008 and 2020, while they do not decrease for median income households. These predictions indicate that the income growth is not the same for everyone.

Figure 6 shows this more clearly by presenting the growth of the 10th, median, and 90th percentile of the income distribution between 1989-2020 for the elderly of age 65-90. Like in the other figures, period effects are excluded for the predictions and the realizations. Pensioners

Figure 6: Indexed growth of equivalized household income for the elderly of age 65-90



Income growth for the 10th percentile, the median and the 90th percentile. The solid lines are realizations corrected for period effects, the dashed lines are predictions made with the most extended version of the microsimulation model (the third specification).

with median household income experience the highest income growth. As a result, inequality (indeed) increases in the lower part of the distribution and decreases in the upper part of the distribution. Relative poverty thus increases. On average the income growth of pensioners is predicted to be higher in the future than it was in the past. When we compare the realized average income growth of pensioners between 1989-2007 with the predicted average income growth between 2007-2020 in figure 6, we find an increase in the average income growth per year for median income households from 0.8% until 2007 to 1.4% after 2007. The average income growth of the 90th percentile also increases, from 0.7% per year until 2007 to 1.0% after 2007. The 10th percentile experiences a decrease in the average growth rate from 0.9% to 2007 to 0.3% after 2007. The results of specification one and two are presented in appendix C and lead to similar conclusions, indicating that the explicit modeling of demographic and labor market changes is not very important for investigating the future income distribution, when using fixed effects and modeling the error terms.

The lower part of the income distribution experiences a relatively low income growth. In this part of the distribution there are many households without occupational pension income. The question arises whether the growing inequality in the lower part of the distribution is caused by an increase in the inequality between households with and without occupational pension income. To answer this question we do a Theil decomposition, concentrating on the lower half of the income distribution. Appendix D describes the Theil decomposition method and table 11 shows the results.

[Table 11]

In the lower half of the income distribution, 21% of the households receive no occupational pension in 2010. In 2020 this proportion will shrink to about 15%. As expected, average income is higher for households with occupational pension income, compared to the households without occupational pension income. The Theil index is about two times higher for households without occupational pension income, but the inequality growth between 2010 and 2020 is higher for the households with occupational pension income. The Theil decomposition shows that in 2010, 11% of the inequality in the lower half of the distribution is caused by the inequality between the group of households with and without occupational pension income. By 2020 this is reduced to 5%. The increased inequality in the lower part of the distribution is thus not caused by a higher inequality between households with and without occupational pension income. Instead, the inequality between these two groups will decrease. This means that inequality between households with occupational pension income on the one hand and inactive/self-employed households without pension arrangements on the other will not increase.

8 Conclusions

This paper simulates the income distribution of the Dutch elderly using an open dynamic microsimulation model with cross-sectional aging. The model takes into account developments in household compositions (e.g. more divorces), developments in labor market states (e.g. higher female participation rates), and productivity differences between cohorts resulting in income differences. Furthermore, we consider differential mortality and increased longevity.

Methodologically the model contributes to the existing literature on microsimulation models by taking into account the persistency and heteroskedasticity of income shocks. This has the advantage that it decreases the need to model all underlying processes influencing income. This can improve the balance of refinement and manageability of microsimulation models. Another advantage is that this increases the possibilities to use administrative data, that are less detailed than surveys but are more representative, to make reliable predictions for the whole population.

To illustrate the inclusion of persistent and heteroskedastic income shocks in a microsimulation model and show the balance between refinement and manageability of a microsimulation model we applied three model specifications, with different levels of refinement. The simulation results are about the same and we find that income shocks are persistent and that persistency increases over the lifecycle (even after the correction for fixed effects). As expected, the variance of income shocks is larger for working-age households than for retirement-age households, and is relatively large for households without labor and/or occupational pension income.

The results indicate that average income increases for future generations of pensioners. More specifically, we find that between 2008 and 2020 household income increases on average by about 0.6% per year in the age group 50-64 and 1.0% for pensioners of age 65-90. Income growth is not the same for everyone. Among pensioners of age 65-90, households with median income experience the highest income growth. During the years 2008-2020 their income is predicted to grow on average by about 1.2% per year, while this is about 1.0% for households at the 90th percentile and only 0.5% for households at the 10th percentile of the income distribution. The trend that pension incomes increase relatively the most for median income households may also hold in other OECD countries, where trends in female labor force participation and longevity were similar.

Inequality indices such as the decile ratio p90/p10 and the Gini coefficient show that inequality among pensioners in the age group 65-90 will probably increase up to 2012 and stabilize thereafter. However, a closer inspection of the whole distribution reveals that inequality will probably grow in the lower part of the distribution, while it declines in the upper part of the distribution. The growing inequality in the lower half of the income distribution is probably not caused by an increasing inequality between households with and without occupational pension income. Instead, inequality between households with and without occupational pension income will probably decrease. The contradictory movements in the lower and upper part of the distribution underline the importance of investigating the whole income distribution by using a microsimulation model, instead of just analyzing the development of one inequality index such as the Gini coefficient.

References

Alessie, R., Camphuis, H., Kapteyn, A., 1990. Een simulatiemodel ter berekening van het sociale risico: eindverslag. Economics Institute Tilburg.

Baltagi, B., 1995. Econometric analysis of panel data. John Wiley and Sons Ltd, Chichester.

Bourguignon, F., 1979. Decomposable income inequality measures. Econometrica 47, 901–920.

Bovenberg, A., Ter Rele, H., 2000. Generational accounts for The Netherlands: an update. International tax and public finance 7 (4), 411–430.

Buhmann, B., Rainwater, L., Schmaus, G., Smeeding, T., 2005. Equivalence scales, well-being, inequality, and poverty: sensitivity estimates across ten countries using the Luxembourg income study (LIS) database. Review of Income and Wealth 34 (2), 115–142.

Butricia, B., Iams, H., Moore, J., Waid, M., 2001. Methods in modeling income in the near term (MINT I). ORES working paper n.91.

Cameron, A., Trivedi, P., 2005. Microeconometrics: methods and applications. Cambridge University Press.

Caminada, C., Goudswaard, K., 2001. International trends in income inequality and social policy. International Tax and Public Finance 8, 395–415.

CBS, 2009. Documentatierapport Inkomenspanel onderzoek (IPO). Centraal Bureau voor de Statistiek, Voorburg.

CBS, 2010. Documentatierapport gemeentelijke basisadministratie (gba) 1995-2010v1. Centraal Bureau voor de Statistiek, Voorburg.

Cowell, F., 1980. On the structure of additive inequality measures. Review of Economic Studies 47, 521–531.

Creedy, J., Duncan, A., 2002. Behavioural microsimulation with labour supply responses. Journal of economic surveys 16 (1).

De Jong, A., Van Huis, M., 2003. Huishoudensprognoses 2002-2050: ontwikkelingen naar burgerlijke staat. Statistics Netherlands, Bevolkingstrends, 2e kwartaal.

De Kam, F., Nijpels, F., 1995. Tijdbom: vergrijzing dreigt de AOW, onze pensioenen en de ouderenzorg onbetaalbaar te maken. Uitgeverij Contact, Amsterdam.

Deaton, A., Paxson, C., 1994. Saving, growth and aging in Taiwan. In: Wise, D. (Ed.), Studies in the Economics of Aging. National Bureau of Economic Research, Ch. 5, pp. 331–362.

Deelen, A., 1995. De ontwikkeling van de inkomensverdeling in Nederland op lange termijn. Onderzoeksmemorandum 121, CPB.

Dekkers, G., Busslei, H., Cozzolino, M., Desmet, R., Geyer, J., Hoffmann, D., Raitano, M., Steiner, V., Tanda, P., Tedeschi, S., Verschueren, F., 2008. A classification and overview of micro simulation models, and the choices made in MIDAS. In: What are the consequenses of the AWG projections for the adequacy of social security pensions. Report of the Work Package 4 of the AIM project.

Emmerson, C., Reed, H., Shephard, A., 2004. An assessment of PENSIM2. IFS working paper WP04/21, London: Institute for fiscal studies.

Euwals, R., Knoef, M., van Vuuren, D., 2011. The trend in female labour force participation: what can be expected for the future? Empirical Economics 40 (3), 729–753.

Flood, L., Klevmarken, A., Mitrut, A., 2006. The income of the Swedish baby boomers. IZA discussion paper, no. 2354.

Geyer, J., Steiner, V., March 2010. Public pensions, changing employment patterns, and the impact of pension reforms across birth cohorts: a microsimulation analysis for Germany. IZA discussion paper, no. 4815.

Gottschalk, P., Smeeding, T., 2000. Empirical evidence on income inequality in industrialized countries. In: Atkinson, A., Bourguignon, F. (Eds.), Handbook of income distribution. Vol. 1. Elsevier, New York, Ch. 5, pp. 262–307.

Gruber, J., Wise, D. A., 2004. Social Security Programs and Retirement around the World: Micro-Estimation. No. grub04-1 in NBER Books. National Bureau of Economic Research, Inc.

Harding, A., 2007. Challenges and opportunities of dynamic microsimulation modelling. Plenary paper presented at the first General Conference of the International Microsimulation Association, Vienna, 21 August 2007.

Haveman, R., Holden, K., Romanov, A., Wolfe, B., 2007. Assessing the maintenance of savings sufficiency over the first decade of retirement. International tax and public finance 14, 481–502.

Hayashi, 2000. Econometrics. Princeton University Press.

Kalmijn, M., Alessie, R., 2008. Life course changes in income: an exploration of age- and stage effects in a 15-year panel in The Netherlands. Netspar Panel Paper 10.

Kalwij, A., Alessie, R., Knoef, M., 2011. Individual income and remaining life expectancy at the statutory retirement age of 65 in The Netherlands. Forthcoming in Demography.

Kapteyn, A., Alessie, R., Lusardi, A., 2005. Explaining the wealth holdings of different cohorts: productivity growth and social security. European Economic Review 49, 1361–1391.

Kapteyn, A., De Vos, K., 1999. Social security and retirement in the netherlands. In: Gruber, J., Wise, D. A. (Eds.), Social security programs and retirement around the world: The relationship to youth employment. Chicago University Press, Ch. 8, pp. 243–259.

Kapteyn, A., De Vos, K., Kalwij, A., 2010. Early retirement and employment of the young in The Netherlands. In: Gruber, J., Wise, D. (Eds.), Social Security Programs and Retirement around the World: The relationship to youth employment. University of Chicago Press, Ch. 8.

Knoef, M., Alessie, R., Kalwij, A., 2009. Changes in the income distribution of the dutch elderly between 1989 and 2020: A microsimulation, netspar discussion paper No. 09/2009-030.

Knoef, M., De Vos, K., 2009. Representativeness in online panels: how far can we reach, paper presented at the MESS workshop 2008, www.lissdata.nl.

Law, A., Kelton, W., 1982. Simulation Modelling and Analysis. New York: McGraw-Hill.

Martini, A., Trivellato, U., 1997. The role of survey data in microsimulation models for social policy analysis. Labour 11 (1), 83–112.

Merz, 1991. Microsimulation - a survey of principles, developments and applications. International Journal of Forecasting 7, 77–104.

O'Donoghue, C., 2001. Dynamic microsimulation: A methodological survey. Brazilian Electronic Journal of Economics 4 (2).

OECD, 2004. Female labor force participation: past trends and main determinants in oecd countries, economics department, working paper.

OECD, 2011. Pensions at a glance 2011: Retirement-income systems in oecd and g20 countries, www.oecd.org/els/social/pensions/PAG.

Osler, M., Prescott, E., Grønbæk, M., Christensen, U., Due, P., Engholm, G., 2002. Income inequality, individual income, and mortality in Danish adults: analysis of pooled data from two cohort studies. British Medical Journal 324, 13–16.

Panis, C., Lillard, L., 1999. Near term model development part II, final report. Santa Monica: Rand Corp.

SCP, 2003. Inkomen verdeeld: Trends in ongelijkheid, herverdeling en dynamiek. Netherlands Institute for Social Research, number 2003/3.

SCP, 2006. Rapportage ouderen 2006; veranderingen in de leefsituatie en levensloop. Netherlands Institute for Social Research, number 2006/12.

Shorrocks, A., 1980. The class of additively decomposable inequality measures. Econometrica 48, 613–625.

Smith, K., Favreault, M., Ratcliffe, C., Butricia, B., Toder, E., 2007. Modeling income in the near term 5. The Urban Institute, Washington D.C.

Soede, A., Vrooman, J., Ferrasi, P., Segre, G., 2004. Unequal welfare states: distributive consequences of population ageing in six european countries. Netherlands Institute for Social Research (SCP) and Center for research on pensions and welfare policies (CeRP).

SZW, 2009. Onderzoek effecten afschaffing AOW-partnertoeslag. Ministry of Social Affairs and Employment.

Toder, E., Uccello, C., O'Hare, J., Favreault, M., Ratcliff, C., Smith, K., Burtless, G., Bosworth, B., 1999. Modeling income in the near term - projections of retirement income through 2020 for the 1931-1960 birth cohorts. The Urban Institute, Washington D.C.

Van Ewijk, C., Draper, N., Ter Rele, H., Westerhout, E., 2006. Aging and the sustainability of Dutch public finances. Special publication 61, CPB.

Von Gaudecker, H., Scholz, R., 2007. Differential mortality by lifetime earnings in Germany. Demographic Research 17 (4), 83–108.

Wolf, D., 2001. The role of microsimulation in longitudinal data analysis. Canadian Studies in Population 28, 165–179.

Zaidi, A., Rake, K., 2001. Dynamic microsimulation models: a review and some lessons for SAGE. SAGE discussion paper 2, Simulating social policy in an ageing society.

A Estimation method

To estimate the parameter vector β of equation (1) we compute

$$y_{it} - \overline{y}_i = \beta(x_{it} - \overline{x}_i) + (v_{it} - \overline{v}_i), \tag{A.1}$$

where $\overline{y}_i$ is the average household income of household i across time. $\overline{x}_i$ and $\overline{v}_i$ are average values across time for each household i. Using (A.1), we can consistently estimate β with OLS. Furthermore, we correct the standard errors of β for the fact that the error terms $v_{it} - \overline{v}_i$

are correlated for observations of the same household (see for example Cameron and Trivedi 2005, p. 727). For the computation of the variance-covariance matrix standard regression routines use $\sum_{i=1}^{N}(T_i) - K$ in the denominator of the multiplier, where T_i denotes the number of periods household i is in the data and K the number of explanatory variables. However, $\sum_{i=1}^{N}(T_i - 1) - K$ should be used here. Therefore, we multiply the variance-covariance matrix of the standard regression routine with $\left(\sum_{i=1}^{N}(T_i) - K\right) / \left(\sum_{i=1}^{N}(T_i - 1) - K\right)$, see for example (Baltagi, 1995), p.12.

Averaging (1) across all observations gives

$$\overline{y} = \alpha + \beta\overline{x} + \overline{v}, \tag{A.2}$$

when we use the identifying assumption $\sum_{i=1}^{N}\mu_i = 0$. From (A.2), we obtain $\widehat{\alpha}$ by

$$\widehat{\alpha} = \overline{y} - \widehat{\beta}\overline{x}, \tag{A.3}$$

We also average the data of individual households across time

$$\overline{y_i} = \alpha + \beta\overline{x_i} + \mu_i + \overline{v_i}, \tag{A.4}$$

from (A.4) we compute $\widehat{\mu_i} = \overline{y_i} - \widehat{\alpha} - \widehat{\beta}\overline{x_i}$. Now $\widehat{v}_{it}$ is computed by

$$\widehat{v}_{it} = y_{it} - \widehat{\alpha} - \widehat{\beta}x_{it} - \widehat{\mu_i}. \tag{A.5}$$

which we use in the auxiliary regression for autocorrelation (3).

B Extended estimation results

Table B.1: Extended estimation results for table 3, the age dummy coefficients

	Coef 1	SE	Coef 2	SE	Coef 3	SE
age 36	-0.259	0.0050	-0.201	0.0051	-0.217	0.0055
age 37	-0.249	0.0049	-0.187	0.0050	-0.204	0.0054
age 38	-0.240	0.0048	-0.177	0.0050	-0.195	0.0054
age 39	-0.227	0.0047	-0.165	0.0049	-0.184	0.0053
age 40	-0.211	0.0047	-0.151	0.0049	-0.173	0.0053
age 41	-0.191	0.0046	-0.138	0.0048	-0.161	0.0052
age 42	-0.170	0.0045	-0.128	0.0047	-0.153	0.0051
age 43	-0.150	0.0045	-0.122	0.0046	-0.147	0.0051
age 44	-0.123	0.0044	-0.109	0.0046	-0.136	0.0050
age 45	-0.099	0.0043	-0.101	0.0045	-0.129	0.0049
age 46	-0.072	0.0043	-0.089	0.0044	-0.118	0.0048
age 47	-0.046	0.0042	-0.078	0.0043	-0.107	0.0048
age 48	-0.019	0.0042	-0.062	0.0042	-0.092	0.0047
age 49	0.006	0.0041	-0.045	0.0041	-0.075	0.0046
age 50	0.020	0.0041	-0.034	0.0040	-0.064	0.0045
age 51	0.031	0.0040	-0.025	0.0039	-0.053	0.0044
age 52	0.043	0.0039	-0.011	0.0039	-0.038	0.0043
age 53	0.046	0.0039	-0.005	0.0038	-0.031	0.0042
age 54	0.051	0.0038	0.005	0.0037	-0.018	0.0041
age 55	0.051	0.0037	0.010	0.0036	-0.011	0.0040
age 56	0.047	0.0036	0.011	0.0035	-0.005	0.0039
age 57	0.038	0.0036	0.008	0.0035	-0.005	0.0038
age 58	0.033	0.0035	0.007	0.0034	-0.002	0.0037
age 59	0.024	0.0034	0.003	0.0033	-0.002	0.0035
age 60	0.017	0.0033	-0.001	0.0032	-0.001	0.0033
age 61	0.008	0.0032	-0.006	0.0031	-0.001	0.0031
age 62	-0.003	0.0030	-0.013	0.0029	-0.005	0.0029
age 63	-0.007	0.0028	-0.013	0.0028	-0.004	0.0028
age 64	-0.008	0.0026	-0.012	0.0025	-0.002	0.0026
age 65	0		0		0	
age 66	-0.015	0.0024	-0.013	0.0023	-0.011	0.0023
age 67	-0.017	0.0026	-0.012	0.0025	-0.010	0.0025
age 68	-0.019	0.0028	-0.012	0.0027	-0.010	0.0026
age 69	-0.021	0.0029	-0.012	0.0028	-0.010	0.0028
age 70	-0.022	0.0030	-0.010	0.0029	-0.009	0.0029
age 71	-0.026	0.0032	-0.012	0.0031	-0.011	0.0030
age 72	-0.025	0.0033	-0.009	0.0032	-0.009	0.0031
age 73	-0.023	0.0033	-0.005	0.0032	-0.004	0.0032
age 74	-0.019	0.0034	0.000	0.0033	0.000	0.0033
age 75	-0.019	0.0036	0.002	0.0035	0.001	0.0034

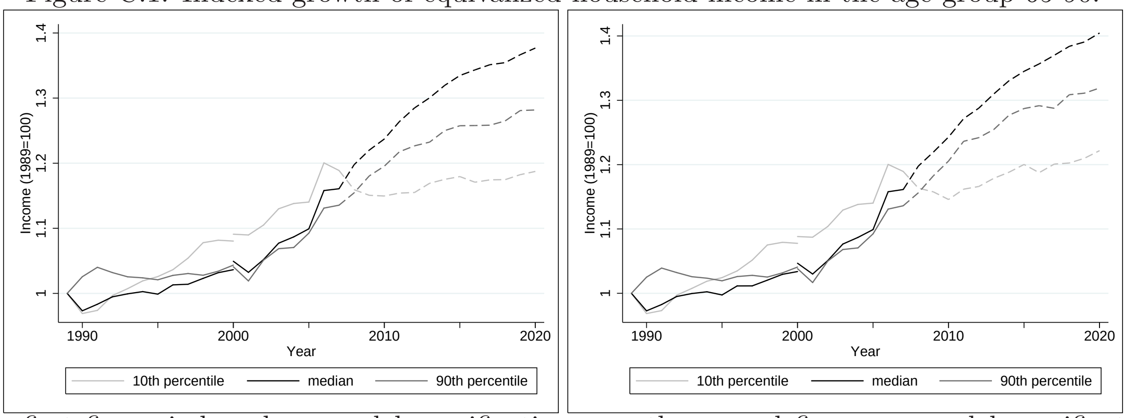
The reference category is 'age 65'. Estimation results continue on the next page.

Table B.1 continued

	Coef 1	SE	Coef 2	SE	Coef 3	SE
age 76	-0.018	0.0037	0.005	0.0036	0.003	0.0036
age 77	-0.016	0.0039	0.008	0.0039	0.007	0.0038
age 78	-0.010	0.0041	0.016	0.0040	0.014	0.0040
age 79	-0.007	0.0043	0.020	0.0043	0.018	0.0042
age 80	-0.001	0.0045	0.028	0.0044	0.024	0.0043
age 81	0.002	0.0047	0.032	0.0047	0.028	0.0046
age 82	0.006	0.0051	0.038	0.0051	0.033	0.0050
age 83	0.004	0.0057	0.039	0.0056	0.033	0.0055
age 84	0.011	0.0062	0.048	0.0061	0.041	0.0060
age 85	0.018	0.0071	0.057	0.0069	0.048	0.0068
age 86	0.035	0.0079	0.075	0.0077	0.066	0.0076
age 87	0.031	0.0098	0.074	0.0096	0.063	0.0095
age 88	0.032	0.0131	0.074	0.0124	0.061	0.0122
age 89	0.050	0.0155	0.092	0.0146	0.081	0.0144
age 90	0.060	0.0222	0.099	0.0207	0.091	0.0211

C Indexed growth in specification 1 and 2

Figure C.1: Indexed growth of equivalized household income in the age group 65-90.



The first figure is based on model specification one, the second figure on model specification two. The dashed lines are predictions and the solid lines are realizations corrected for period effects.

D Theil decomposition

Overall income inequality can be related to mutually exclusive population subgroups using a Theil decomposition. We use the Theil decomposition method to explore whether the rising inequality in the lower part of the distribution is caused by an increase in the inequality between households with and without occupational pension income. Theil decompositions are developed by Shorrocks (1980), Bourguignon (1979), and Cowell (1980). The Theil index is a weighted average of inequality within subgroups, plus inequality among those subgroups. More specific, inequality within a year is the average inequality within each subgroup, weighted by the income of the subgroups, plus the inequality among subgroups. The subgroups in this study are (1) households with occupational pension income and (2) households without occupational pension income. The Theil index is given by

$$T = \frac{1}{N}\sum_{i=1}^{N}\frac{y_i}{\overline{y}}\log\left(\frac{y_i}{\overline{y}}\right) \tag{D.1}$$

where N is the number of observations, y_i the income of household i, and $\overline{y}$ average income of all households. (D.1) can be rewritten as

$$T = (s_1 T_1 + s_2 T_2) + \left(s_1 \log\left(\frac{\overline{y_1}}{\overline{y}}\right) + s_2 \log\left(\frac{\overline{y_2}}{\overline{y}}\right)\right) \tag{D.2}$$

where the first term presents within group inequality and the second term presents between group inequality. T_1 is the Theil index for households with occupational pension income, T_2 is the Theil index for households without occupational pension income. s_1 is the share of total income received by the households with occupational pension income, s_2 is the share of total income received by the households without occupational pension income. $\overline{y_1}$ is the average income of households with occupational pension income and $\overline{y_2}$ is the average income of households without occupational pension income.

Tables

Table 1: Transitions in marital status[a]

Year	Unmarried → Married	Married → Divorced	Widow → Married	Divorced → Married
1995	1.42	0.68	0.38	2.67
1996	1.46	0.67	0.49	2.66
1997	1.57	0.67	0.56	2.77
1998	1.59	0.67	0.45	2.70
1999	1.59	0.71	0.45	2.78
2000	1.59	0.80	0.40	2.49
2001	1.72	0.79	0.41	2.53
2002	1.66	0.78	0.38	2.45
2003	1.60	0.77	0.39	2.21
2004	1.61	0.81	0.37	2.14
2005	1.56	0.81	0.37	2.25
2006	1.58	0.81	0.38	2.26

[a] Percentage of observations making a transition in marital status. For example, of all unmarried persons in 1995, 1.42% make a transition into marriage. Source: GBA, selection of all persons in the age group 36-90 with year of birth between 1917 and 1970. Own computations.

Table 2: Descriptives equivalized household income[a]

Year	Mean	p10	p50	p90	$\frac{p90}{p10}$	$\frac{p90}{p50}$	$\frac{p50}{p10}$	Gini
Age 50-64								
1989	20114	11310	18346	30705	2.71	1.67	1.62	0.228
1992	21183	11473	19242	32495	2.83	1.69	1.68	0.241
1995	21718	11320	19490	34049	3.01	1.75	1.72	0.250
1998	22747	12025	20534	35206	2.93	1.71	1.71	0.246
2001	24203	12838	21786	37468	2.92	1.72	1.70	0.247
2004	24463	13124	22035	37641	2.87	1.71	1.68	0.245
2007	24351	12814	21528	38257	2.99	1.78	1.68	0.258
Age 65-90								
1989	17031	10355	14699	26732	2.58	1.82	1.42	0.225
1992	17626	10542	14935	28176	2.67	1.89	1.42	0.236
1995	17278	10605	14659	27246	2.57	1.86	1.38	0.228
1998	17916	11275	15192	27758	2.46	1.83	1.35	0.221
2001	18562	11702	15737	28252	2.41	1.80	1.34	0.224
2004	19189	12073	16366	29316	2.43	1.79	1.36	0.222
2007	20048	12406	17196	30592	2.47	1.78	1.39	0.227

[a] Source: IPO, own computations.

Table 3: Estimation results fixed effects income equation

	Coef 1	SE	Coef 2	SE	Coef 3	SE
age dummies[a]	yes		yes		yes	
year dummies	yes		yes		yes	
# adult men			0.131	0.0018	0.037	0.0028
# adult women			0.061	0.0018	-0.030	0.0023
# children			-0.068	0.0015	-0.059	0.0015
widower			0.138	0.0072	0.084	0.0071
widow			0.044	0.0051	-0.044	0.0058
divorced (man)			0.033	0.0062	0.021	0.0060
divorced (woman)			-0.123	0.0077	-0.140	0.0076
unmarried (man)			0.057	0.0091	0.050	0.0088
unmarried (woman)			-0.071	0.0120	-0.080	0.0119
# labor (man)					0.120	0.0026
# labor (woman)					0.118	0.0018
# occ. pension (man)					0.058	0.0032
# occ. pension (woman)					0.099	0.0034
$\rho_{0,1}$	-0.074	0.6553	0.789	0.6698	0.984	0.6705
$\rho_{1,1}$ (age/10)	-0.464	0.4741	-1.050	0.4844	-1.182	0.4848
$\rho_{2,1}$ ((age/10)^2)	0.303	0.1261	0.438	0.1288	0.468	0.1289
$\rho_{3,1}$ ((age/10)^3)	-0.052	0.0146	-0.065	0.0149	-0.068	0.0149
$\rho_{4,1}$ ((age/10)^4)	0.003	0.0006	0.003	0.0006	0.003	0.0006
$\rho_{5,1}$ (age>65)	0.116	0.0099	0.122	0.0101	0.126	0.0101
$\rho_{6,1}$ (age=63)	0.034	0.0082	0.039	0.0083	0.044	0.0083
$\rho_{7,1}$ (age=64)	0.070	0.0086	0.080	0.0087	0.082	0.0087
$\rho_{8,1}$ (age=65)	-0.123	0.0089	-0.121	0.0089	-0.128	0.0089
$\rho_{9,1}$ (age=66)	-0.184	0.0094	-0.195	0.0096	-0.201	0.0096
ρ_2	0.065	0.0012	0.054	0.0012	0.055	0.0012
α	9.909		9.746		9.805	
σ_μ	0.370		0.369		0.342	
σ_ϵ	0.210		0.205		0.202	
R^{2}[b]	0.6881		0.7122		0.7225	
N	861336		861336		861336	

Reference categories are ‘age 65’ and ‘married’. For the identification of age, period, and cohort effects the method of Deaton and Paxson (1994) is used. Clustered standard errors are used to take into account the correlation of the error terms in the same household.

[a] The coefficients of the age specific dummy variables can be found in Appendix B

[b] These R^2’s take into account the household-specific effects.

Table 4: Logit transition models for marital status

	Men		Women	
	Coef	SE	Coef	SE
Married → Divorced				
age/10	1.461	0.0414	1.611	0.0482
$(\text{age}/10)^2$	-0.190	0.0043	-0.224	0.0052
(year of birth-1900)/10	0.179	0.0404	-0.097	0.0477
$((\text{year of birth-1900})/10)^2$	0.013	0.0037	0.038	0.0042
constant	-8.556	0.0894	-7.888	0.1054
pseudo R^2	0.050		0.056	
Divorced → Married				
age/10	-0.388	0.0560	0.167	0.0704
$(\text{age}/10)^2$	-0.032	0.0057	-0.098	0.0074
(year of birth-1900)/10	-0.068	0.0539	0.290	0.0708
$((\text{year of birth-1900})/10)^2$	-0.010	0.0050	-0.031	0.0064
constant	-0.069	0.1197	-2.975	0.1533
pseudo R^2	0.028		0.050	
Unmarried → Married				
age/10	-2.732	0.0751	-1.954	0.1066
$(\text{age}/10)^2$	0.203	0.0086	0.124	0.0123
(year of birth-1900)/10	1.654	0.0919	1.927	0.1305
$((\text{year of birth-1900})/10)^2$	-0.138	0.0077	-0.151	0.0108
constant	-0.918	0.1987	-4.154	0.2755
pseudo R^2	0.046		0.065	
Widowed → Married				
age/10	1.254	0.1353	1.234	0.1467
$(\text{age}/10)^2$	-0.175	0.0114	-0.188	0.0130
(year of birth-1900)/10	0.204	0.0919	0.419	0.1101
$((\text{year of birth-1900})/10)^2$	-0.026	0.0106	-0.018	0.0118
constant	-5.406	0.3319	-7.481	0.3376
pseudo R^2	0.092		0.143	

Table 5: Logit transition models for children

Children not leaving their parental home	Coef	SE
age child	-0.008	0.0147
age child^2	-0.005	0.0004
male	0.474	0.0445
constant	4.072	0.1198
pseudo R^2	0.165	
N	38906	
New children being born	**Coef**	**SE**
age/10[a]	-5.983	0.1449
(age/10)^2	0.459	0.0154
(year of birth-1900)/10	-0.432	0.1666
((year of birth-1900)/10)^2	0.073	0.0144
man	0.679	0.0232
couple	1.514	0.0473
one child	0.753	0.0287
two children	-0.637	0.0306
constant	11.220	0.4473
pseudo R^2	0.219	
N	850151	

[a] Age of the key person.

Table 6: Logit transition models for the labor market status of single men and women

	Labor → Occup. pension				Labor → Other			
	Men		Women		Men		Women	
	Coef	SE	Coef	SE	Coef	SE	Coef	SE
age/10	0.234	0.6802	0.121	0.7283	-1.194	0.2637	-1.001	0.2838
$(age/10)^2$	0.091	0.0457	0.020	0.0476	0.114	0.0218	0.090	0.0227
(year of birth-1900)/10	-0.139	0.2471	-0.812	0.2743	-0.314	0.0634	-0.341	0.0733
interaction age and cohort	-0.020	0.0457	0.116	0.0488	-0.009	0.0206	-0.041	0.0218
divorced	0.240	0.0802	-0.414	0.0800	0.303	0.0538	0.454	0.0592
widow(er)	0.345	0.1096	0.797	0.0805	-0.077	0.1227	0.048	0.1065
# children			-0.099	0.0659			0.137	0.0202
constant	-6.315	2.3818	-3.256	2.6181	1.601	0.6178	1.954	0.7104
pseudo R^2	0.063		0.104					
N	127759		83337					

	Other → Labor				Other → Occup. pension			
	Men		Women		Men		Women	
	Coef	SE	Coef	SE	Coef	SE	Coef	SE
age/10	-0.760	0.2741	-0.486	0.2744	2.067	0.8009	0.680	0.6978
$(age/10)^2$	0.017	0.0227	-0.017	0.0222	-0.133	0.0513	-0.090	0.0426
(year of birth-1900)/10	0.139	0.0644	0.254	0.0684	0.205	0.3114	-0.882	0.2963
interaction age and cohort	-0.045	0.0219	-0.065	0.0218	-0.046	0.0524	0.139	0.0456
divorced	0.502	0.0590	0.582	0.0579	0.318	0.0907	-0.222	0.0795
widow(er)	0.219	0.1410	0.441	0.1036	0.642	0.1074	0.960	0.0718
# children			-0.014	0.0169			-0.113	0.0546
constant	0.803	0.6192	-0.164	0.6538	-10.647	2.9932	-3.676	2.7678
pseudo R^2	0.183		0.245					
N	40821		49339					

Estimation results of the transition models for labor market status. We assume 'occupational pension' to be an absorbing state. The interaction between age and cohort is formally 'age/10*(year of birth-1900)/10'.

Table 7: Multinomial logit model for the initial labor market status of new household members and children entering adulthood

	Men		Women	
	Coef	SE	Coef	SE
Labor				
age/10	8.778	0.0404	6.150	0.0371
$(\text{age}/10)^2$	-1.768	0.0105	-1.134	0.0095
$(\text{age}/10)^3$	0.107	0.0008	0.063	0.0008
(year of birth-1900)/10	0.207	0.0700	-1.925	0.0657
$(\text{year of birth-1900/10})^2$	0.074	0.0127	0.560	0.0115
$(\text{year of birth-1900/10})^3$	-0.007	0.0007	-0.036	0.0006
constant	-13.502	0.1349	-10.038	0.1377
Occupational pension				
age/10	-4.883	0.1267	0.545	0.1512
$(\text{age}/10)^2$	1.548	0.0221	0.248	0.0251
$(\text{age}/10)^3$	-0.104	0.0013	-0.020	0.0014
(year of birth-1900)/10	0.834	0.0690	0.064	0.0623
$(\text{year of birth-1900/10})^2$	-0.062	0.0157	0.076	0.0158
$(\text{year of birth-1900/10})^3$	-0.001	0.0012	-0.005	0.0013
constant	-5.856	0.2926	-10.375	0.3422
pseudo R^2	0.489		0.296	
N	1019127		998296	

We distinguish three labor market states: labor, occupational pension, and 'other' (no labor and no occupational pension income). 'Other' is the reference category in this estimation.

Table 8: Predictions of marital status[a]

	Men				Women			
Year	Marr	Unmarr	Wid	Div	Marr	Unmarr	Wid	Div
Age 50-64								
2008	75.6	9.9	2.0	12.5	71.8	7.5	6.2	14.5
2011	72.4	12.1	2.0	13.4	69.4	9.0	5.8	15.8
2014	68.9	14.5	1.9	14.7	67.4	10.7	5.1	16.9
2017	64.9	17.3	1.8	15.9	65.0	12.5	4.7	17.7
2020	61.2	20.1	2.0	16.8	61.9	15.2	4.5	18.4
Age 65-90								
2008	74.5	5.6	12.4	7.5	46.4	5.7	39.5	8.4
2011	73.2	5.8	12.3	8.7	48.4	5.5	36.6	9.5
2014	72.3	5.9	11.9	10.0	50.2	5.3	34.1	10.5
2017	70.8	6.4	11.9	10.9	50.4	5.5	32.4	11.8
2020	68.8	7.4	12.0	11.8	50.0	5.7	31.0	13.3

[a] Marital status for men and women. E.g. in 2020 about 61.2% of men in the age group 50-64 will be married.

Table 9: Predictions of labor market status[a]

	Men			Women		
Year	Labor	Occupational pension	Other	Labor	Occupational pension	Other
Age 50-64						
2008	62.6	19.6	17.8	46.3	15.0	38.6
2011	62.5	23.7	13.8	50.0	19.2	30.8
2014	64.2	24.3	11.5	54.2	22.3	23.5
2017	64.5	26.0	9.5	57.2	24.5	18.3
2020	66.0	25.6	8.4	59.4	26.8	13.8
Age 65-90						
2008	3.6	87.0	9.4	2.1	54.0	43.8
2011	3.7	87.8	8.5	2.5	56.2	41.3
2014	4.4	88.4	7.2	3.2	59.7	37.1
2017	4.1	89.6	6.2	3.0	65.4	31.5
2020	4.4	90.6	5.0	3.2	71.1	25.7

[a] In case a person receives both labor income and occupational pension income the labor market status is based on the highest income component. E.g. in 2020 labor is the most important income source for 66.0% of men in the age group 50-64.

Table 10: Predictions of income[a]

Year	Mean	p10	p50	p90	$\frac{p90}{p10}$	$\frac{p90}{p50}$	$\frac{p50}{p10}$	Gini
Age 50-64								
2008	24559	13325	22040	37995	2.85	1.72	1.65	0.243
2011	24996	13795	22541	38357	2.78	1.70	1.63	0.241
2014	25531	14009	22989	39244	2.80	1.71	1.64	0.239
2017	25690	13923	23081	39754	2.86	1.72	1.66	0.241
2020	26173	13855	23303	41102	2.97	1.76	1.68	0.244
Age 65-90								
2008	20285	12245	17837	31296	2.56	1.75	1.46	0.225
2011	21227	12378	18936	32828	2.65	1.73	1.53	0.229
2014	21962	12641	19670	33661	2.66	1.71	1.56	0.228
2017	22445	12841	20163	34236	2.67	1.70	1.57	0.230
2020	22905	13048	20725	35073	2.69	1.69	1.59	0.230

[a] In this paper income is always inflated/deflated to 2005 euro's. This table shows the results of the most extended model specification, where demographic variables and labor market states are taken into account (model specification three).

Table 11: Theil decomposition of equivalized household income

Year	2010	2015	2020
% Households without occupational pension	20	18	15
Average income, households without occ. pension	12455	13358	14008
Average income, households with occ. pension	14880	15705	16059
Theil index, households without occ. pension	0.035	0.036	0.039
Theil index, households with occ. pension	0.014	0.016	0.022
Within group inequality	0.0174	0.0196	0.024
Between group inequality	0.0024	0.0018	0.0011
% Between group inequality	12	9	4

[a] This table concentrates on the lower half of the income distribution of pensioners (age 65-90). It shows the inequality within and between households with and without occupational pension income.