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

Marike Knoef † Rob Alessie ‡ Adriaan Kalwij §

February 2010

*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 (CBS) for their help in tracking down the changes in the IPO data due to the revision in 2000. We thank Koen Caminada, Didier Fouarge, Hans-Martin von Gaudecker, Theo Nijman, Arthur van Soest, participants of EALE 2009, and participants of the conference 'Labor Force Participation and the Well-being of the 50+ population' (2009) in Utrecht for their useful comments.

†Tilburg University, Netspar, and CentERdata

‡University of Groningen, Netspar, Tinbergen Institute, and CentER

§Utrecht University, Netspar, and IZA

## Abstract

Ageing 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. To achieve a sustainable development of public budgets, policy reforms are needed. Therefore, policymakers require insights into the income distribution of current and future retirees. The main objective of this paper is to quantify these distributions. First, this paper describes the income distribution of retirees in the past. Secondly, we estimate the development in the income distribution until 2020, due to increased longevity and developments in households' demographic and socio-economic status. Examples of these are the increased number of divorces and increased female labour force participation. We use a microsimulation model that takes into account differential mortality, allows for parallel and serial interactions between characteristics and individuals, takes into account that the distribution of income shocks may be different for different types of households, and allows the persistency of income shocks to vary with age.

Results indicate that average income increases for future generations of retirees. Between 2008 and 2020 equivalised household income will increase on average about 0.6% per year for the age group 50-64 and 1.0% for retirees (65-90). Among retirees, the highest predicted annual growth is for median-income households (1.2%). High-income households have a somewhat lower growth (1.0%) and low-income households only have a predicted yearly income growth of 0.5%, which yields implications for inequality. Overall indices such as the decile ratio p90/p10 and the Gini coefficient show that for the retirees income inequality rises up to 2012 and stabilizes thereafter. The upward trend up to 2012 is due to an increasing inequality in the lower part of the distribution, while inequality declines in the upper part of the distribution. A Theil decomposition shows that the inequality increase in the lower part of the distribution is not a result of an increasing inequality between households with and without occupational pension income. Instead, inequality between these households decreases.

*Keywords*: income distribution, population ageing, microsimulation.
*JEL codes*: J14, I3, D3.

## 1 Introduction

In 2011 the first generation of the babyboom will reach the statutory retirement age of 65. From then onwards, there will be a doubling in the proportion of retirees over the working population from 26% in 2011 to 47% in 2038. This places an increasing financial burden on society through pay-as-you-go financed social security, pension, health, and long-term care systems.

Policies aimed at alleviating the costs related to the ageing society can be based on the notion of intergenerational solidarity such 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. An example is the 'fiscalisation' of the public pension contributions. In this case, a larger part of the pay-as-you-go public pension scheme will be financed by general tax revenues. Consequently, also the 65+ population pays for the state pensions and due to the progressivity of the Dutch tax system, this policy option will result in a redistribution of income within the elderly generation.

In order to assess the viability of proposed reforms, policymakers require insights into the income distribution of current and future retirees in the situation of no policy changes. One should realize that also without pension reforms the future income distribution of retirees will be different from the current one, because of developments in longevity, demographic and socio-economic compositions. For instance, the number of divorces has increased, widowhood among women tends to decrease as a result of the life expectancy of men converging to that of women, and female labor force participation has increased tremendously during the last decades, which leads to more women receiving occupational pension income in the future. The increased female labour force participation has changed the distribution of work between households into more 'all-work' and 'no-work' households. In the future this may lead to more 'all-occupational pension' and 'no occupational pension' households.

The main objective of this paper is first to describe the income structure of retirees in the past and secondly to predict the income structure for next generations of retirees. For these purposes this paper exploits administrative panel data containing very detailed information on individual and household income of a representative sample of the Dutch population over the period 1989-2007. First, we describe developments in the income distribution of the age groups 50-64 and 65-90 between 1989-2007. We also present developments in the income composition for different parts of the distribution. This informs us about the degree to which different parts of the distribution are involved in reforms with regard to state pensions or occupational pensions. An interesting finding is that occupational pensions have become more important over the whole income distribution (not just at the upper part).

Secondly, we predict the income distribution of the elderly until 2020 by microsimulation, using an open dynamic population model with cross-sectional ageing. The advantage of the microsimulation approach is that detailed distributional estimates are possible. The results can for example be used in the development of tax modifications to pay the costs of the ageing society.

In the model differential mortality risks are taken into account, such that relatively poor households have a relatively high probability to decease and relatively rich households have lower probabilities to decease. For the income predictions we estimate an income equation using a fixed effects model with three specifications: the first specification only contains age and period effects, in the second specification household demographics are added, and in the third specification also the labour market status of household members are taken into account. The fixed effects allow for permanent differences among households due to unmeasured variables. Since we want to derive information about the income distribution, we explicitly pay attention to the modeling of the error terms. We take autocorrelation into account and allow for the fact that the degree of autocorrelation differs over the lifecycle. Furthermore, we allow the distribution of income shocks to be different for younger and older households and for households where members participate in the labour market and/or receive an occupational pension income or not. We are not aware of a previous microsimulation study on income that takes into account heteroskedasticity and persistency of income shocks, and where the persistency of an income shock is allowed to depend on age. The results show that the next generations of retirees contain less widows, due to the converging life expectancy of men and women. In addition, the proportion of women with occupational pension income increases about 16.4%-points between 2008 and 2020. With regard to income we find that next generations of retirees have higher equivalised household incomes than current generations of retirees, especially among households with median income. Between 2008 and 2020 equivalised household income of retirees increases on average 0.5% per year for the 10th percentile, 1.2% for the median and 1.0% for the 90the percentile. Inequality among retirees increases at the lower part of the income distribution, but decreases at the upper part of the distribution. The increased inequality in the lower part of the distribution is not the result of a higher inequality between households with an without occupational pension income. Instead, inequality between households with and without occupational pension income decreases until 2020.

Previous research making predictions about the income distribution in the Netherlands is Dessens and Jansen (1997). They examined the consequences of the increased proportion of female-partners going to work on trends in income inequality. They extrapolated the Gini coefficient until

2011, using predictions of female participation rates and the average ratio of their incomes to those of their partners. By contrast, we examine developments in the whole income distribution. Our empirical results indicate that a single inequality index such as the Gini coefficient, is not always informative enough to describe trends, because trends in the lower and upper part of the distribution may be contradictive.

SZW (2006) also predicts the income distribution of future retirees in the Netherlands. They do use a microsimulation approach, such that detailed estimates on the whole income distribution are possible. However, their baseline data is a survey on housing needs in 2002[1] of which the representativeness is questionable. The response rate of this survey is 60%, with large differences between different types of households, types of houses, and types of neighborhoods. In addition, item nonresponse in known to be particularly severe for sensitive questions about income, the variable of interest. In a microsimulation model the quality of the input data is of main importance: when the baseline data are not representative, the predictions of the population will not be representative either (Martini and Trivellato, 1997). Our study is based on administrative data, where nonresponse does not play a role. In addition we use panel data, such that we are able to measure flows and to model individual change, which is not possible in the cross sectional data used by SZW (2006). In SZW (2006) income components are simulated separately, which gives more information than our aggregated income measure, but is sensitive to correlations between the developments of the several income components. For example, a high education level may lead to a relatively high wage rate, but also to a relative high rate of return on capital, which should be taken into account. A more recent microsimulation model in The Netherlands is presented by Van Sonsbeek (2009). This model is focused on predictions of the costs of state pensions in the future.

With regard to other countries, Flood et al. (2006) simulate the income of Swedish babyboomers. They have panel data and estimate individual earning profiles with a random effects model. In the simulation random numbers, independent of the covariates, are drawn to represent the random effects. It is thus very important that the individual effects are orthogonal to the covariates of the model. Satisfying this assumption is difficult. For example, motherhood may be part of the individual effect, as it may influence the number of hours worked (and therewith earnings). However, it is also correlated with the covariate martial status. Instead, we use a fixed effects model that does not suffer from this problem. The fixed effects approach is possible because our target population are retirees and the income profiles are estimated with the same data that form the basis of the population model. In several other countries dynamic microsimulation models are used for pension issues. Dekkers et al. (2008) present the model MIDAS for

[1]Woning Behoeften Onderzoek 2002

Belgium, Italy, and Germany and provide an overview on relevant dynamic microsimulation models used in the US, UK, and Italy.

If ones aim is to quantify the effects of different pension policies, than it is important to model labour supply responses explicitly (Creedy and Duncan, 2002). This is out of the scope of this paper. This paper gives insights into the development of the future income distribution, induced by increased longevity and ongoing demographic and socio-economic changes. If labour 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 reviews the empirical literature on the distribution of income of the elderly in the Netherlands. Sections 3 and 4 describe the data and the microsimulation model after which in section 5 the estimation results are summarized. Section 6 presents the results of the simulation and finally, section 7 concludes.

# 2 Literature on the income position of the elderly in the Netherlands

This section provides a short review of the empirical literature on relevant income trends in the Netherlands. We discuss the following topics: 1) the income distribution in the Netherlands and the position of the elderly therein, 2) income over the lifecycle, and 3) developments in the income composition of the elderly.

## 2.1 Income distribution and the position of the elderly therein

Before paying attention to the income distribution of the elderly, it is important to know something about the development of the total income distribution of the last decades. This section therefore discusses empirical literature to find out how and through which pathways the total distribution of income has changed, and what the position of the elderly in this distribution has been.

As from the 1960s inequality had decreased rapidly. Reason for this was the relative increase of the income of inactive households, induced by the construction of the social security system (Trimp, 2000; Caminada and Goudswaard, 2003). Later, as from 1979 inequality started to increase. In a cross-country analysis, Gottschalk and Smeeding (2000) found that, although the Netherlands began from a relative low base Gini, income inequality in the Netherlands has increased relatively fast between 1979-1994. Caminada and Goudswaard (2001) state that the two main forces behind this phenomenon are a more unequal distribution of market incomes and changes in social transfers. In 1990 a revision of the tax system lead to more inequality. In addition, the strong increase of double income households changed the income distribution. SCP (2003) found that the increased number of two earner couples has increased inequality between 1985 en 1994. Using a decomposition analysis, they find that the increase in two earner couples explains in combination with the relative decreasing incomes of people without labour income, about one third of the total increase in inequality. The increase in the number of two earner couples can also lead to a pooling effect: the inequality within the group of households with two earners is lower than that of households with one earner. Therefore, as from a certain point, an increase in the proportion of two earner households reduces household income inequality. In the future, the trend of more two earner households will result in more households receiving two occupational pension incomes.

In the second half of the 90s income inequality between households has decreased slightly (De Vos, 2007), whereas income inequality between household appears to be quite stable during 2000 to 2007 (CBS statline). With regard to the position of the elderly in the income distribution, De Vos (2007) found that households in the age group 50-59 are overrepresented in the highest deciles and that this overrepresentation has increased between 1989 and 2000. Furthermore, people between 60-64 are underrepresented in the lowest deciles in 1989. To a lesser extent this also holds in 2000. Another interesting result is that the relative income position of the 65+ households per equivalent adult has not improved between 1989 and 2000, despite the fact that supplementary occupational pensions have increased. De Vos mentions the increased labour force participation of women as a possible explanation for the relative improvement in the financial position of people younger than 65.

## 2.2 Income over the lifecycle

Income at retirement is related to income earlier in life. For that reason this section describes some literature about income during the life course. On average, households whose head is between 55 and 64 years old, have the highest disposable income SCP (2006). In addition, generation effects are important for income. Kapteyn, Alessie, and Lusardi (2005) found that productivity growth can explain all generation effects. De Vries and Kalmijn (2007) and Kalmijn and Alessie (2008) have related life course changes in income per equivalent adult to the role of partnership and parenthood transitions. They find that there is a positive partner-effect on the income position, especially for women. Further, they do not find any negative effect of widowhood for both men and women, but parenthood transitions are very strong. This is partly due to a decline in personal income of woman after the birth of children, but mainly because of an increase in the expenditures which are necessary for children.

### 2.3 Income composition of the elderly

Income consists of several components. Households with a head between 55-64 mostly receive wage income, early retirement pension, unemployment benefits, disability benefits or other social security benefits. At age 65 the main sources of income are state pensions and occupational pensions. SCP (2006) reports that between 1994 and 2003 the share of earnings in total household income has increased for elderly between age 55-64. At the same time, income sources as early retirement have become less important, probably because early retirement schemes have become less generous to discourage early exits from the labour force. Between 1994 and 2003, the income share of state pension has decreased at the expense of occupational pensions: whereas basic state pension provided 62% of the income in 1989, it was reduced to 54% in 2003 (De Vos, 2007). The income share of occupational pensions, on the other hand, increased from 30% in 1989 to 40% of the income of the 65+ in 2003. This finding might be partly explained by spending cuts of the government during the early nineties, such that public pension benefits have not been adjusted for inflation.[2] Another explanation for the increase of the income share of occupational pensions is the development of the pension system in the 50's and the 60's, which made a higher rate of people receiving occupational pension (Deelen, 1995). Furthermore, higher participation rates lead to more people receiving occupational pensions (when reaching retirement age). As from 2015 the partner bonus, a bonus for the younger partners of state pension beneficiaries with no or a low income, will be abolished. This will lower state pension income for households who do not become eligible anymore as from 2015. However, remaining household income for most of these households will not reach the limit for social assistance (SZW, 2009).

## 3 Data

The data used in this paper are from the Dutch Income Panel (Inkomens Panel Onderzoek, IPO) and from the population register (Gemeentelijke Basis Administratie, GBA). Statistics Netherlands has gathered these data. Section 3.1 provides basic information about IPO and descriptives on income and labour market status. Section 3.2 describes the information we use from the population register.

[2]In principle, state pensions follow the gross minimum wages, which are linked to the development of the contractual wages. In case no inflation correction takes place, public pensions lag behind the growing prosperity. All the more because contractual wages itself lag behind earned incomes, because of occasional increments and promotions. (De Kam and Nijpels, 1995)

### 3.1 Dutch income panel (IPO)

IPO contains information about households and their income, based on administrative data. Most of these data are from the Dutch National Tax office. Additional data are from registrations of rent subsidies and subsidies for the financing of study. In the Dutch Income Panel, so called 'key persons' are randomly drawn from the Dutch population and are followed over time. Also information on all household members of the key persons are available. Because of the administrative character, we have a very low attrition rate, which is a very nice feature for this research. In surveys it is well-known that the rich and the poor are often under represented and that institutional households are not included (Alessie et al., 1990). Another advantage of administrative data is that the observed variables are measured with high accuracy. A drawback of the Dutch Income Panel is that it lacks some crucial background variables, such as education levels. Variables which 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.

For our model many waves of a panel are necessary to disentangle age, period, and cohort effects. We have waves from 1989 to 2007 at our disposal, which means that we have data over 19 years. In 2000 a revision in the Dutch Income Panel has taken place. During this revision several data sources, definitions, and methods have been changed. Appendix A mentions several important changes. For the year 2000 two datasets are available: one with the data sources, definitions, and methods before revision and one with the data sources, definitions, and methods after revision. We have tried to equalize the definitions as much as possible. Appendix A provides the steps we have token to accomplish this.

In order to be able to compare the income of households with different compositions and size we use the CBS equivalence scale (see Siermann et al. 2004, for more details). There exists a wide range of equivalence scales. We choose the CBS scale because it is based on the Dutch situation. It takes into account the number of adults, the number of young children, the age of the oldest child, and the age of the person in the household who earns the largest part of the income in the household. Kalmijn and Alessie (2008) found that the modified OECD scale and the CBS scale yield very similar results with regard to the distribution of equivalised household income. Appendix B explains the definition of income. Income in this paper is always deflated/inflated to the prices of 2005 (real income), using the CPI.

Table 7 reports on the selection of the data. At first, we exclude households with missing or non-positive household income[3]. 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. Now we are left with 70% of the raw data. We are interested in households with key persons of age 50-90 during 1989-2020. However, to be able to simulate income of a household with key person of age 50 in 2020 we need income of all households with key persons as from age 36 in 2006. Figure 9 gives an overview of these generations. We select all households where the key person is born between 1917-1970 and is of age 36-90. [4] Finally, households in the bottom or top 0.1% of the income distribution are excluded.

#### 3.1.1 Descriptives household income

Tables 8 and 9 describe equivalised household income, the main variable of this research, for key persons in the agegroups 50-64 and 65-90, respectively. As from now, when we write 'income' we always mean 'equivalised household income'. During the years 1989-2007 income has increased. In the age group 50-64 mean income has increased 21%, from 20,114 in 1989 to 24,351 in 2007. In the age group 65-90, income has been fairly constant during the nineties. It has increased only 1% between 1990-1999, compared to 9% between 2000-2007.[5] The Gini coefficient and the decile ratios show that inequality in the agegroup 50-64 has increased between 1989 and 1995 and was rather constant afterwards.[6] For the agegroup 65-90 inequality is lower and shows a different pattern. It was increasing between 1989-1991, but has decreased in the years after 1991. As from 1998 inequality in the agegroup 65-90 is quite stable. These developments add to the results of Gottschalk and Smeeding (2000), who found that overall income inequality has increased from 1979 to the mid nineties. Several factors may have induced these trends, such as the increased female labour supply (and therewith the increased share of two-earner couples), changes in early retirement schemes, the development of the pension system, and the business cycle.

[3]In 64% of the households with negative income there are one or more self-employed household members. However, of all households with self-employment, only 3,3% have a negative income, such that we do not overlook a large part of the self-employed households

[4]In this study we ignore new immigrant families. For the agegroup under consideration (50-90) we expect the effect of this ignorance to be small.

[5]Probably this is related to the fact that in the nineties no indexation of public pension benefits has taken place.

[6]In 1990 a major revision of the tax system has taken place with distributional consequences ('operatie Oort'). This explains (part) of the difference between inequality measures from 1989 to 1990.

### 3.1.2 Income composition 1989-2007

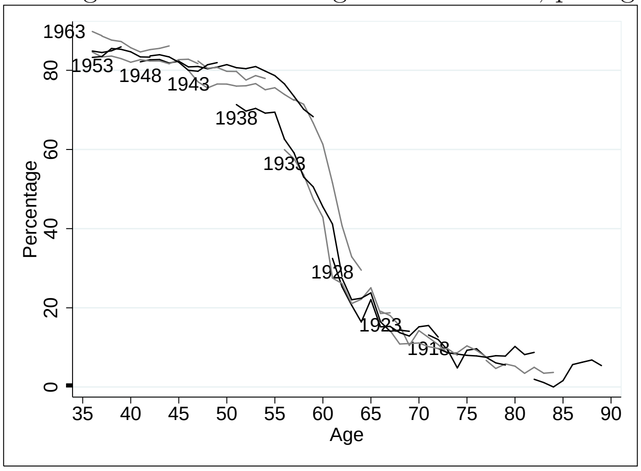
This paper deals with the income distribution of household income, which is the sum of several income components. This section investigates the composition of these income components. Figures 6 and 7 show the income composition in 1989 and 2007 for the age groups 50-64 and 65-90. The horizontal axes give the percentiles of the income distribution, the vertical axes give the proportions of the several income components. For example, in Figure 7 the income of the average household in the 50th percentile (with median income) consists of 32% occupational pension income, 58% public pension benefits, and 10% remaining income sources. For age 50-64 (Figure 6) we see that the proportion of labour income increases over the income distribution: the more income households receive, the more labour is an important component. As expected, for the lower percentiles transfer income is important. Interesting is that in 2007 less households depend on transfer income than in 1989. While in 1989 as from the 25th percentile labour income becomes more important than transfer income, in 2007 this intersection already takes place at the 13th percentile.

With regard to age 65-90 (Figure 7), state pensions and occupational pensions are the most important income sources. Figure 7 indicates to what degree different parts of the distribution are involved in reforms with regard to state or occupational pensions, respectively. Between 1989 and 2007, occupational pensions have become more important for almost all percentiles. In 1989, as from about the 85th percentile occupational pensions were more important then state pensions. In 2007 already as from the 70th percentile occupational pensions are more important than state pensions. This result sharpens the result already found by De Vos (2007) (section 2.3). Above age 65, it are especially the high percentiles who receive labour income. Probably, especially the households with the 'better' jobs remain participating in the labour market and can generate a relative high share of their income out of it. Capital income only plays a substantial role in the top 5% of the households.

### 3.1.3 Labour and occupational pensions

Section 3.1.2 showed that labour and occupational pensions belong to the most important income components. Between 1989 and 2007 labour force participation of both women and men has changed. Obviously this has implications for the income structure of the next generation of retirees, since more labour income today generally leads to more occupational pension income in the future. Figure 1 shows for several generations the proportion of male key persons receiving labour income. For example, '1938' refers to the key persons born in 1938. The vertical differences between lines measure the 'cohort-time' effects. We use this terminology to emphasize that it is not

Figure 1: Percentage of males receiving labour income, per age and cohort




possible to disentangle age from cohort and time effects in this figure. From age 50 to 65 there is a steep decrease in the proportion of males receiving labour income. However, as from the generation born in 1943, more men of age 50-65 keep participating in the labour market. At age 60 about 46% of men born in 1938 received labour income. For the generation born in 1943 this was 61%, so there is a cohort-period effect of 15%-points. This reflects the fact that early retirement schemes have become less generous to discourage early exits from the labour force. These results are in line with Kapteyn et al. (2009), who found that the labour force participation for men of age 55-64 decreased until 1993 and has increased afterwards. Note that the figure only shows the proportion of males receiving labour income. It does not tell anything about the amount of labour.

Figure 2: Percentage of females receiving labour income, per age and cohort

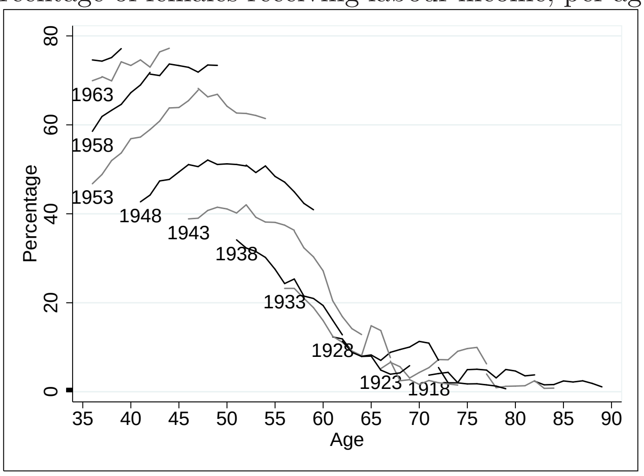


Figure 2 also shows the proportion of female key persons receiving labour income. Female participation rates have increased considerably and it can be seen that there are important cohort-period effects. This is also found by Euwals et al. (2007). They have decomposed the growth in the female labour force participation and found that changed attitudes towards the combination of paid work and children have played a major role. This trend will have considerable consequences for the income structure of next generations of retirees, as more two earner couples today will lead to more couples receiving double pension incomes in the future. In addition, the position of single old women may improve, because of an improvement of their own pension incomes.

In couples retirement decisions are interrelated. The age difference between men and women is on average 2.5 years and is quite stable over time. Figure 8 shows the proportion of women, younger then 65, who have a partner of age 65 or older, and are receiving labour income (compared to the total group of women younger than 65, having a partner of age 65 or older). During the period 1989-2007 this proportion has increased from 11.7% to almost 22.5%. This is an interesting trend, which leads to elderly households receiving more labour income and more occupational pension income in the future.

We already found that occupational pensions have become more important (section 2.3). Table 1 shows that especially the percentage of women receiving an occupational pension has increased. In 1989, 29% of the women

Table 1: Descriptives occupational pension income

| Age | % Pension men | | | % Pension women | | |
|---|---|---|---|---|---|---|
| | 1989 | 1999 | 2007 | 1989 | 1999 | 2007 |
| 36 | 1.42 | 1.09 | | 1.03 | 1.34 | |
| 40 | 1.63 | 1.90 | 2.45 | 1.50 | 1.27 | 2.56 |
| 45 | 1.74 | 2.66 | 3.14 | 2.77 | 3.08 | 4.45 |
| 50 | 2.00 | 5.18 | 3.52 | 5.89 | 6.88 | 5.13 |
| 55 | 8.16 | 8.20 | 9.20 | 10.28 | 8.19 | 10.11 |
| 60 | 34.47 | 33.92 | 31.45 | 19.92 | 18.45 | 28.68 |
| 65 | 84.76 | 88.75 | 96.21 | 28.94 | 36.00 | 62.71 |
| 70 | 84.21 | 89.69 | 91.94 | 41.13 | 40.83 | 46.53 |
| 75 | | 86.70 | 90.22 | | 40.48 | 51.47 |
| 80 | | 85.91 | 87.19 | | 57.03 | 61.96 |
| 85 | | | 83.70 | | | 69.19 |
| 90 | | | 88.89 | | | 67.47 |

This table shows the percentage of men and women with occupational pension income in the years 1989, 1999 and 2007. As we have selected the birth years 1917-1970 there are no descriptives for age 36 in 2007, for age 75+ in 1989, and age 85+ in 1999. Source: IPO, own computations.

of age 65 received an occupational pension, in 2007 this was almost 63%. Over the lifecycle an increasing percentage of women receive an occupational pension income. For example, in 2007 47% of the 70 year old women received an occupational pension income, while 69% of the 85 year old women received an occupational pension income. This increase can be attributed to widow pensions, which women receive when their partners die. New generations of retired men also receive occupational pension income more often. The percentage of men receiving an occupational pension at age 65 has increased from 85% in 1989 to 96% in 2007.

### 3.2 Population register (GBA)

Trends in marital status may influence the income distribution of the next generations of retirees. Divorces, for example, lead to changes in income as well as household formation (and thus the equivalence scale). GBA contains information on marital status of all people registered in Dutch municipalities.[7] Because GBA is a much bigger dataset than IPO, we use GBA to estimate transitions in marital status from one year to another. Data is available from January 1 1995 to January 1 2008. Just as in the Dutch Income Panel we select all persons born between 1917-1970 with age 36-90. Furthermore, to be able to estimate transitions between $t$ and $t+1$, 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 $t$.

Given this selection above, we end up with 6,812,340 individuals in 1995 up to 8,673,138 individuals in 2006 (of whom marital status in $t$+1 is known). Table 10 shows the percentage of people making a transition from one state to another. The percentage of married people who divorce between $t$ and $t+1$ has increased from 0.68% in 1995 to 0.81% in 2006. Furthermore, the table shows that 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, 0.42% of the widows and widowers make a transition into marriage.

## 4 Microsimulation model

This section describes the functioning of the model we use to simulate the income distribution of the elderly until the year 2020. We use a dynamic population model with cross-sectional ageing.[8] Cross-sectional ageing allows us to have interactions between household members. For example, labour

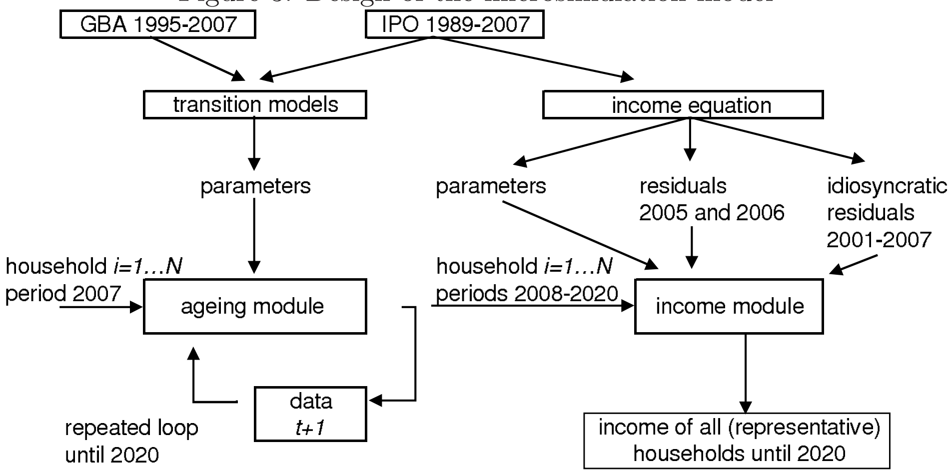
[7] Individuals not registered as residents are, for instance, NATO personal, diplomats and individuals illegaly residing in The Netherlands.

[8] See Dekkers et al. (2008) for a description of the classification of micro simulation models.

market positions may be determined simultaneously, and the death of a household member can influence the behaviour of the remaining household members. The model is open, as marriage and birth lead to new synthetic household members. Figure 3 describes the design of the model. In a nutshell, we dynamically age all members of the representative households in the Dutch Income panel until 2020 and predict their household income. In the ageing module, household members may decease, divorces may take place, children may move out of their parental home, new partners or children may enter the household and labour market positions may change. Transition models are used to estimate these transitions in household composition, marital status, and labour market status. Secondly, we estimate income until 2020 in the income module. Therefore, we estimate an income equation, taking into account age and period effects, household demographics, labour market status of household members, and fixed effects. The fixed effects allow for permanent differences among households due to unmeasured variables. This is in line with Haveman et al. (2007), who found that pre-retirement economic advantages continue into retirement. The fixed effects also include productivity differences that result in income differences between cohorts (Kapteyn et al., 2005). For every household of which we want to predict income in 2008-2020, we know the fixed effect and we add the estimated effects of changed household demographics, labour market status, and the changed position in the lifecycle. To take into account autocorrelation and heteroskedasticity, we use residuals and idiosyncratic residuals from the income equation, respectively.

In the remainder of this section we explain the income equation (4.1) and the transition models with regard to household demographics (4.2) and labour market status (4.3). Finally, the simulation method is described (4.4).

Figure 3: Design of the microsimulation model



## 4.1 Income equation

To predict income trends for future generations of retirees we model household income using a fixed effects model with several explanatory variables, such as the age of the key person of the household, the business cycle, marital status, and the labour market position of household members. The fixed effects give us the opportunity to control for time-invariant omitted variables, such as cohort effects, that influence the income of a household. 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 income of household $i$ in time period $t$, $\alpha$ is a scalar, $x_{it}$ is the $it$-th observation on $K$ explanatory variables, $\mu_i$ is the unobserved individual effect and $v_{it}$ is the error term. We assume strict exogeneity

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

and identify $\alpha$ by assuming $\sum_{i=1}^{N} \mu_i = 0$. The estimation of $\alpha$, $\beta$ and $\mu_i$ is explained in Appendix C.[9]

We estimate three specifications of the income equation (1). In the first specification the vector $x_{it}$ only contains age and period effects. This is the pure specification where mobility only results from income shocks. By adding extra variables to the vector $x_{it}$, more individual heterogeneity in the income path is introduced. In the second specification, demographic variables such as household size and marital status are added, and in the third specification also the labour market position of household members is taken into account. Extra individual heterogeneity is thus introduced by mutations in marital status and labour market positions, which are explicitly modeled and depend on age and cohort. In the more detailed specifications, income shocks thus have more motivation, as part of them are linked to demographic factors and labour market status.

Age and period effects are implemented as dummy variables, such that their relationship with income is very flexible. However, empirically, age, period, and cohort effects (which are captured in the individual effect) cannot be identified, since calendar time is equal to the year of birth plus age. We follow the identification restriction proposed by Deaton and Paxson (1994), which means that we assume that all time dummy coefficients add up to zero and are orthogonal to a linear time trend. We assume that all period effects are due to unanticipated business cycle shocks.

[9] We prefer a fixed effects model to a random effects models as household specific effects ($\mu_i$) may be correlated with included covariates. For example, 'ability' is likely to be correlated with labour market status. In microsimulations, the disadvantage of a fixed effects estimator is that it rules out out-of-sample simulations (Wolf, 2001). In this analysis we can use a fixed effects model because our target population are retirees and income profiles are estimated with the same data as the base population is coming from.

To gather information on the whole income distribution, modeling the residual term $v_{it}$ is very important. Households perceive 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, and may be higher for singles than for couples. Furthermore, the question arises how long these shocks persist (autocorrelation), and whether the persistency of a shock depends on the position in the lifecycle. For example, at older ages, shocks may be more persistent.

We start with the persistency of income shocks: 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[10]

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

where we assume $\epsilon_{it}$ to be serially uncorrelated. The persistency of a shock may depend on the position in the lifecycle.[11] Therefore, we allow $\rho_{it,1}$ to be a function of age.[12]

$$\rho_{it,1} = \rho_{0,1} + \rho_{1,1}(age_{it}/10) + \rho_{2,1}(age_{it}/10)^2 \tag{4}$$

As explained above, the variance of an income shock may also depend on the characteristics of a household. We take this heterogeneity into account by investigating the distribution of $\epsilon_{it}$ for several mutually exclusive groups and distinguish the income shocks of these groups in the simulation (for example, key person is younger or older than 65, and whether household receives labour income). For each group we draw income shocks from the empirical distribution of residuals in 2001-2007 for that group ($\hat{\epsilon}_{i,2001}, ..., \hat{\epsilon}_{i,2007}$).

## 4.2 Transition models marital status and children

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. Transition probabilities between the various states are modeled using multinomial logit models. The transition models are estimated for men and women separately and use age and year of birth as explanatory variables. 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. We assume people to make only one transition in marital status per year.

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

[11] Kalmijn and Alessie (2008) find that the 2-year autocorrelation of standardized income is quite stable during midlife, but moves to a higher level after age 65.

[12] We have tried several specifications and have also investigated whether it is relevant to specify $\rho_2$ as a function of age.

The number of children in a household influences the equivalence factor, and therefore also equivalised household income. The probability of a child moving out of the household is estimated with a logit model, where age and gender of the child are explanatory variables. The probability of a newborn child is also modeled with a logit model. We assume the probability of a newborn child to depend on the age and gender of the key person, whether there is a couple in the household and the number of children which are already present in the household. In reality, also older children may enter the household. In this simulation model we ignore children already born to enter the household.

### 4.3 Labour market status

In the third specification of the income equation three labour market positions are distinguished, (1) receiving labour income, (2) receiving occupational pension income and (3) receiving none of these two ('other'). In order to belong to (1) or (2), labour income or occupational pension income has to be at least 500 euro per year. In case an individual receives both labour income and pension income the highest income component counts.

The transition probabilities are estimated with multinomial logit models for men and women. The explanatory variables used in this estimation are age, cohort, marital status, and the number of children. We assume'occupational pension' to be an absorbing state. The increased labour market participation of women enters the model in two ways: first via the initial labour market positions of women, secondly via cohort effects in the labour market transition models.

The labour market position of two members of a couple are interrelated, both the decision to participate in the labour market and the retirement decision. We therefore estimate transition models for couples and singles separately. We threat the trivariate outcomes of person A and B part of the same couple as 3*3=9 univariate outcomes. Transitions between these 9 states are estimated with a multinomial logit model. Also here we assume 'occupational pension' to be an absorbing state for each member of the couple.

To determine labour market status at time $t + 1$ with the transition models, we have to know the labour market status at time $t$. A problem arises for new household members and children who just enter adulthood. To determine an initial state for them we estimate a simple multinomial logit model per gender, with age and cohort as explanatory variables.

### 4.4 Simulation method

The starting point of the simulation is the collection of all (representative) households in the Dutch Income Panel 2007. We dynamically age all household members, with transitions in household characteristics and labour market states. To determine whether an individual in the sample deceases we draw for each individual $j$ in household $i$ in each period $t$ from 2007 to 2019 a random value $m_{ijt}$ from the uniform distribution. If $m_{ijt}$ is lower than the predicted mortality rate, belonging to the age, cohort, and gender of individual $i$, the individual deceases and is not present anymore in the next period (see e.g. Law and Kelton, 1982). 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. (2009) in the Netherlands. When 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. As from age 65, Kalwij et al. find a quartile ratio Q1/Q4 of 2.2 for men and 1.7 for women, meaning that mortality rates in the first quartile are 2.2 times higher for men and 1.7 times higher for women, relative to the fourth quartile.[13] As from age 65 we therefore adjust mortality rates such that mortality rates in the fourth quartile are 2.2 (or 1.7 for women) times higher as in the first quartile, keeping the average mortality rate equal. Before age 65 mortality rates are small, such that differential mortality will not make a relevant difference. To determine which households belong to the first and the fourth quartile, we make use of the fixed effects estimation of specification 1. The fixed effects of this estimation give us a measure of the 'lifetime income position' of households, as in this specification a correction has been made only for the age profile, the business cycle, and for income shocks.

For transitions in marital status we use the transition models and estimate the transition probabilities for each key person in the microsimulation model, given their age and marital status in period $t$. Furthermore, we draw a random value $k_{it}$ from the uniform distribution. A transition in marital status takes place when $k_{it}$ is smaller than the estimated transition probability. 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. This new household member has the same age as his/her partner and the opposite gender.

Transitions with regard to children and labour market positions are made in the same way. Using the parameters of the transition models, we estimate the transition probabilities for all persons, given their age, marital status, and labour market position in period $t$. Again, random draws from the uniform distribution determine whether a transition takes place.

After determining mortality and the transitions in household demograph-

[13] This is in line with findings in other European countries, e.g. Von Gaudecker and Scholz (2007) for Germany and Osler et al. (2002) for Denmark.

ics and labour market status we project household income using (1), (3), and (4). Income is predicted by

$$\hat{y}_{i\tau} = \hat{\alpha} + \hat{\beta} x_{i\tau} + \hat{\mu}_i + \hat{v}_{i\tau}, \tag{5}$$

for $\tau$=2008, ..., 2020. $\hat{v}_{i,\tau}$'s are computed using $\hat{v}_{i,\tau-2}$, $\hat{v}_{i,\tau-1}$, and random draws $\epsilon_{i,\tau}$. For example, $\hat{v}_{i,2008}$ is computed by

$$\hat{v}_{i,2008} = (\hat{\rho}_{0,1}+\hat{\rho}_{1,1}(age_{i,2008}/10)+\hat{\rho}_{2,1}(age_{i,2008}/10)^2)\hat{v}_{i,2007}+\hat{\rho}_2\hat{v}_{i,2006}+\epsilon_{i,2008}, \tag{6}$$

$\epsilon_{i,\tau}$ are drawn from the empirical distribution of the residuals from 2001 to 2007.[14] We investigate whether the distribution of the residuals is different for households with different characteristics and will draw $\epsilon_{i,\tau}$ from these sub-distributions. Finally, as from 2015, we take into account the abolishment of the partner bonus for state pensions (described at the end of section 2.3). For all households who are not eligible for a partner bonus anymore, and of whom the younger member of the couple has no labour income, we subtract the partnerbonus from household income.

# 5 Estimation results

This section discusses the estimation results of the income equation explained in section 4.1. The estimation results of the transition models described in section 4.2 and 4.3 are discussed in Appendix D. Table 11 shows the estimation results. The first two columns of the table show the estimation results of the first specification, where only age, period effects, and fixed effects are taken into account. In the second specification household demographics are added and in the third specification also labour market positions are taken into account.

In all three specifications age effects increase until about age 55 and decrease afterwards. As from age 70 they increase again, even when we control for selectivity by adding selection dummies. The shape of the age profiles of specification 2 and 3 are very similar, while the age profile of specification 1 is more pronounced. The estimated period effects follow the development of the business cycle. The persistency of income shocks are estimated by the $\rho$'s. It appears that income shocks are indeed persistent and increase with age. In the first specification $\rho_1$ ranges from 0.29 at age 36 to 0.50 at age 80. In the second and third specification $\rho_1$ is a little bit smaller until age 70. The added demographic variables can capture part of the effect. $\sigma_\mu$ and $\sigma_\epsilon$ show that the individual variation is larger than the random component.

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

In specification two, where household demographics are added to the estimation, it appears that households with more adults have on average a higher equivalised household income. On average they thus bring more income than 'costs' (in terms of the equivalence scale). Households with more children, on the other hand, have a lower income (children cost money in terms of the equivalence scale). Marital status of the key person significantly influences income. Compared to households where the key person is married, divorced men are relatively well off and divorced women are relatively worse off. However, note that a divorce often coincides with a loss of an adult in the household, such that the total effect for men in this case is a 2.8% loss of income (0.033-0.061) and for women a 25% loss of income (-0.123-0.131). Unmarried men and women are better off then divorced men and women. Widowers and widows are relatively well off financially. One has to keep in mind that all coefficients are based on *within* variation.

Specification three adds variables with regard to the labour market status of household members, namely the number of men receiving labour income, the number of women receiving labour income, the number of men receiving occupational pension income, and the number of women receiving occupational pension income. These variables can be considered endogenous explanatory variables, but although this implies the model estimates cannot be given a causal interpretation, they can be used for predicting income. By adding these extra variables, the coefficients of the household demographics change. The estimation results show that men are financially better off than women and that this difference is even larger when persons are widowed, divorced or unmarried (compared to married). As expected, working men and women increase household income and men and women with an occupational pension increase household income in a smaller degree.

Future income shocks are drawn from the empirical distribution of the idiosyncratic residuals in 2001-2007. As expected, the distribution of these residuals is different for different groups of households.[15] We distinguish households with key persons younger and older than 65 and find that the standard deviation of the residual is 40% higher for households where the key person is younger than 65. In the third specification we also distinguish households who receive labour or occupational pension income or 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 labour or occupations pension income, compared to households with labour 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. For the sim-

[15] Kalmijn and Alessie (2008) found that the variance of equivalised income (logged) is relatively low after age 65.

ulations these results implicate that we incorporate higher income shocks for younger households and for households without labour and/or occupational pension income.

## 6 Simulation results

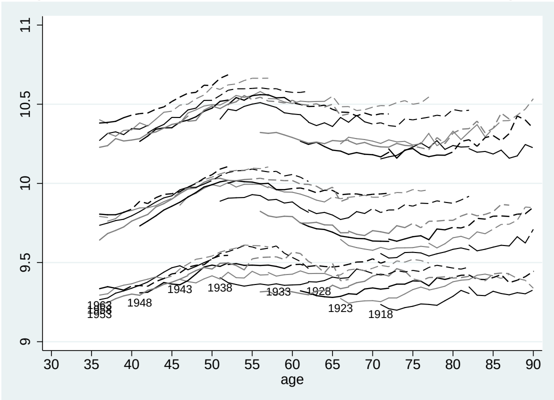
Corresponding to the three specifications of the income equation, we have three future predictions of the income distribution. The third specification represents the most extensive prediction, where also household demographics and labour market positions are taken into account. Before explaining the income predictions we will describe the predictions of marital status and labour market status, as they are input for the income predictions. Predictions of marital status for the years 2008-2020 are given in Table 12 and 13 for the age groups 50-64 and 65-90, respectively. In the age group 50-64, the proportions of married men and women decrease over time, while more men and women become unmarried or divorced. The share of divorced men increases from 13 to 17%, while that of women increases from 15 to 19%. In the age group 65-90 most striking is the decrease in widowhood for women. This has to do with the converging life expectancies of men and women, which leads to younger cohorts of women being widowed less often. Furthermore, the fall in widowhood for women can be attributed to the babyboom generation reaching age 65. Therefore, the total age group 65-90 starts to contain relatively a lot of 'young' elderly who are widowed less often.

Table 14 and 15 give predictions of labour market status for all key persons between 2008-2020. 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 their tremendous increase in labour force participation.

Using the predictions of marital status and labour market status described above, we predict equivalised household income for all households in the three specifications. Table 16 and 17 show the results for the agegroups 50-64 and 65-90, respectively. Incomes in these tables are free from period effects, such as the effects of the business cycle. According to the predictions, income will increases on average 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 increase until about 2012 and stabilizes thereafter. Focussing on the decile ratios p90/p50 and p50/p10, it appears that there are two contradictory developments going on: an increasing inequality in the lower part of the income distribution and a decreasing inequality in the upper part of the income distribution. Here, the importance of investigating the whole income distribution, instead of just an inequality measure such as the Gini coefficient, becomes clear. Inequality indices differ in their sensitivities to income differences in different parts of the distribution, but one index can not show the different developments going on in the whole income distribution. For the age group 50-64 the Gini coefficient and the decile ratio p90/p10 show that inequality decreases a bit until 2015, but increases a bit afterwards.

Figure 4 shows realizations and predictions of log income per age and cohort. Period effects are excluded for the predictions, as well as for the realizations. On purpose, we show the 10th, 50th, and 90th percentile of *log* equivalised household income, as it is more interesting to compare relative than absolute changes. The age profile of the 50th and 90th percentile is

Figure 4: Log equivalised household income per age and cohort.



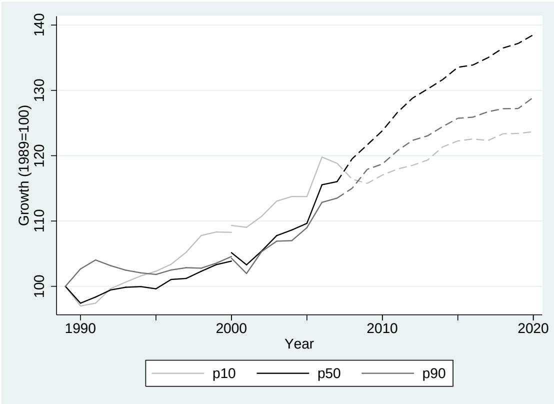

This figure shows the 10th, median, and 90th percentile of income per age and cohort. The dashed lines are predictions, solid lines are realizations corrected for period effects.

relatively strong, compared to that of the 10th percentile. Probably, in the lowest income group less variety over the lifecycle is possible. As expected, younger cohorts have higher incomes than older cohorts. However, between 2008 and 2020 cohort-time effects decrease for the 10th percentile while they do not decrease for the 50th percentile. It becomes clear that the income growth is not the same for everyone.

To show this more thoroughly, Figure 5 presents the growth of the 10th, 50th, and 90th percentile of the income distribution between 1989-2020 for retirees. Also in this figure period effects are excluded, to make a good comparison. The results of specification 1 and 2 are presented in Figure 10 and are almost the same.

Retirees with median household income experience the highest income growth, such that inequality (indeed) increases in the lower part of the distribution and decreases in the upper part of the distribution. Relative poverty thus increases. When we compare the realized average income growth of retirees between 1989-2007 with the predicted average income growth between 2007-2020, we find a decrease in the average income growth per year

Figure 5: Indexed growth of equivalised household income for retirees (age 65-90)




Dashed lines are predictions, solid lines are realizations corrected for period effects.

for the 10th percentile from 0.9% until 2007 to 0.3% after 2007. The average income growth of the median increases, from 0.8% per year until 2007 to 1.4% after 2007. The 90th percentile also experiences an increase in the average growth rate from 0.7% to 2007 to 1% after 2007. The income growth of retirees will be higher in the future than it was in the past, especially for the median. However, the lower part of the income distribution experiences a lower growth. In this part of the distribution there are relatively a lot of households without occupational pension.

The question arises 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. To explore this question we use a Theil decomposition analysis, concentrating on the lower half of the income distribution. Appendix E describes the Theil decomposition method. The results can be found in Table 2. In the lower half of the income distribution 21% of the households receives no occupational pension in 2010. In 2020 this proportion has decreased to 15%. As expected, average income is higher for households with occupational pension, compared to the households without occupational pension income. The Theil index is about two times higher for households without occupational pension income, but the increase in inequality between 2010 and 2020 is higher for the households with occupational pension income. In the decomposition we find 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. In 2020 this is reduced to 5%. The increased inequality in the lower part of the distribution is thus not caused by a higher between group 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 on the one hand, and inactive/self-employed households without pension arrangements on the other hand will not increase.

Table 2: Theil decomposition

| Year | 2010 | 2015 | 2020 |
|---|---|---|---|
| % Households without occ pension | 21 | 18 | 15 |
| Average income, households without occ. pension | 12608 | 13448 | 13859 |
| Average income, households with occ. pension | 14825 | 15776 | 16030 |
| Theil index, households without occ. pension | 0.033 | 0.039 | 0.039 |
| Theil index, households with occ. pension | 0.013 | 0.016 | 0.022 |
| Within group inequality | 0.0167 | 0.0197 | 0.0240 |
| Between group inequality | 0.0020 | 0.0017 | 0.0012 |
| % Between group inequality | 11 | 8 | 5 |

This table concentrates on the lower half of the distribution and shows the inequality within and between households with and without occupational pension income.

## 7 Conclusions

This paper examines the income distribution of the elderly between 1989-2007 and predicts income of the elderly between 2008-2020, using an open dynamic microsimulation model with cross-sectional ageing.

Exploration of the data reveals that between 1989-2007, equivalised household income of the elderly in the age group 50-64 has increased on average 1.1% per year. Inequality in this age group has risen between 1989-1995 and was rather stable afterwards. The income of the elderly in the age group 65-90 has remained fairly constant during the nineties and has increased on average 1.3% per year between 2000-2007. Income inequality has increased between 1989-1991, but decreased thereafter. Occupational pensions have become more relevant for the whole income distribution between 1989 and 2007.

As discussed in the introduction, one should be aware that without any policy intervention (e.g. pension reforms) the income distribution of future retirees will already be different from the current one. Using a microsimulation model, we find out how the income distribution develops as a results of changes in longevity, demographic and socio-economic compositions (such as the increased number of divorces, the increased female labour force participation and increased productivity which lead to changes in income).

Concerning demographic transitions, we find that the percentage of widows in the age group 65-90 will decrease by 9%-points until 2020. In addition to a composition effect, this result can be attributed to the fact that life expectancies of men and women converge. With regard to labour market status, especially the proportion of women reveiving an occupational pension income increases. We predict that the proportion of women with an occupational pension (including widow pensions) increases from 54 in 2008 to 70% in 2020.

For the prediction of the income distribution of future generations of retirees we have estimated an income equation using a fixed effects model. Furthermore, we have taken into account differential mortality and explicitly modeled the distribution of the error terms over time and across households. We found that income shocks are persistent and that persistency increases over the lifecycle. The dispersion of an income shock is larger for 'young' households and for households without labour and/or pension income.

The results of the microsimulation model indicate that average income increases for future generations of retirees. To be more specific, we find that between 2008 and 2020 household income increases on average with about 0.6% per year in the age group 50-64 and 1.0% for retirees (65-90). Income growth is not the same for everybody. Among retirees, households with median income experience the highest income growth. During the years 2008-2020 their income is predicted to grow on average 1.2% per year, while this is 1.0% for the 90th percentile and only 0.5% for the 10th percentile.

Inequality indices such as the decile ratio p90/p10 and the Gini coefficient show that inequality among retirees rises up to 2012 and stabilizes thereafter. However, a closer inspection of the whole distribution reveals that this upward trend up to 2012 is due to an increasing inequality in the lower part of the distribution, while in the upper part of the distribution inequality decreases. The rising inequality in the lower half of the income distribution is not caused by an increasing inequality between households with and without occupational pension income. Instead, inequality between households with and without occupational pension income is decreasing. It are households with small occupational pensions for whom the income growth is relatively low. The contradictive movements in the lower and upper part of the distribution underline the importance of investigating the whole income distribution, here achieved by using microsimulation, instead of just analyzing the development of one inequality index such as Gini coefficient.

The results suggests that a policy such as fiscalisation as discussed in the introduction can be effective to combat the financial burden of population ageing because a majority of the future retirees will be considerably wealthier than the current ones. Obviously, this policy measure will further reduce income inequality of the elderly.

## References

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

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.

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.

Caminada, C., Goudswaard, K., 2003. Verdeelde zekerheid. De verdeling van baten en lasten van sociale zekerheid en pensioen. Sdu Uitgevers.

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 Kam, F., Nijpels, F., 1995. Tijdbom. Uitgeverij Contact.

De Vos, K., 2007. 'remote access' final report. Netspar pilot study.

De Vries, J., Kalmijn, M., 2007. Life course changes in income: The role of partnership and parenthood transitions. Mimeo, Tilburg University.

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.

Dessens, J., Jansen, W., 1997. Future trends in household income inequality in the netherlands. The Netherlands journal of social sciences 33 (2), 113–129.

Euwals, R., Knoef, M., Van Vuuren, D., 2007. The trend in female labour force participation: what can be expected for the future? IZA discussion paper 3225.

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

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

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.

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., 2009. Individual income and remaining life expectancy at the statutory retirement age of 65 in the netherlands. Working paper.

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., Kalwij, A., 2009. 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.

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.

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.

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

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

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

Siermann, C., Van Teeffelen, P., Urlings, L., 2004. Equivalentiefactoren. Statistics Netherlands.

SZW, 2006. De toekomstige inkomenspositie van ouderen. Ministry of Social Affairs and Employment/SZW.

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

Trimp, L., 2000. Inkomensverdeling 19591998. Centraal bureau voor de Statistiek, Voorburg/Heerlen, Ch. 5, pp. 41–53.

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

Van Sonsbeek, J., 2009. Micro simulations on the effects of ageing-related policy measures: The social affairs department of the netherlands ageing and pensions model. SSRN working paper.

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.

# A IPO before and after the revision of 2000

In this appendix we number some important changes which took place during the revision of IPO. In addition we mention several steps we have taken to equivalise the data before and after revision as much as possible.

- As from 2000 also one-off income such as severance pays are included. (All income that previously belonged to 'bijzonder tarief', large amounts that do not occur frequently).
- As from 2000 new data sources have been used. In particular with regard to rents and dividends. Before 2000 we did not observe rents and dividend for the people who were only obliged to pay tax on wages (rents belonging to the tax free allowance). As from 2000 new datasources are used such that rents and dividends are observed for everyone (small amounts of income for a rather large group of observations). In order to smooth the data before and after revision we have imputed the rents of 2000 to the years before revision, taking inflation into account.
- Computations of the rental value of real estate are revised.

- Income on individual level has limitations, because certain components (such as child benefits and rent subsidies) are ascribed to a different household member after revision (for example breadwinner instead of the head of the household).

- The method to determine whether persons on a certain address constitute a household together is changed (for example, two students living on the same address).

- Employer contributions from wages are included in the wages before revision but are separated afterwards. To equivalise the definition we subtract them before revision.

- A same kind of thing happens with transfer income. Contributions to social insurance, paid by the authority who pays out transfer incomes were included before revision, but were separated after revision. Unfortunately is was not possible to subtract these contributions before revision, therefore we added them after revision to equivalise the definition before and after revision.

- After revision dividends from stocks of a substantial holding[16] also includes the sellings of stocks from own business. Before revision these were excluded. We try to exclude these dividends by dropping dividends which exceed 250,000 euro.

- In addition to the above mentioned changes, several other changes have taken place, such as the equivalence scale and the categories of various variables.

- Although we have tried to make definitions as consistent as possible, differences are left. Therefore we smooth the years before revision such that they are in line with the years after revision. Smoothing takes place on the individual level. We have used absolute differences and also here we take inflation into account. The income components which we smooth separately are: labour income, occupational pensions, public pension benefits, and remaining transfer income (e.g. disability benefits and unemployment benefits)

- Sometimes the birth year of individuals change over time. For these people we have imputed the birth year from the central municipal basis administration (GBA), which is available as from 1995. There are a few observations who have inconsistent birth years and are not present in GBA (for example because they died before 1995). These observations got a missing birthyear.

[16] A taxpayer is regarded as having a substantial holding in a corporation if he or she, either alone or with his or her spouse, holds directly or indirectly 5% of the issued capital.

# B Construction of net noncapital and net capital income

This appendix describes how we construct net noncapital and net capital income in IPO. We make a distinction between the data before 2001 and after 2001, as in 2001 a new tax system was introduced. At first we will define and explain net noncapital and net capital income between 1989 and 2000, after that we will define and explain net noncapital and net capital income as from 2001.

Net noncapital income between 1989 and 2000 is defined by

$$\text{net noncapital income} = L + T - \frac{L+T}{L+T+H+C}\tau_i - P + \text{allowances},$$

where $L$ is the sum of all income obtained with labour, $T$ is the sum of all transfer income, $C$ is the sum of all capital income, $\tau_i$ is the total taxation on income (from labour, transfers, interests etc.), and $P$ is the sum of the forced premia for social security insurance and employees' insurance. In the Dutch law mortgage interests are tax-deductible. Furthermore, the law states that home owners earn a taxable income from an owner-occupied house (the so called 'imputed rent'). The imputed rent is a percentage of the value of the house determined by the municipal authority. In our calculations of the net capital and noncapital income we take this deductible mortgage interests and imputed rent into account. $H$ therefore is the imputed rent minus the mortgage interests. Allowances consist of child benefits, rent subsidies etc.

For the period 1989-2000 net capital income is defined by

$$\text{net capital income} = (H - \frac{H}{L+T+H+C}\tau_i) + (C - \frac{C}{L+T+H+C}\tau_i - \tau_w),$$

where $\tau_w$ is the tax on wealth. Net capital income consists of two parts. The first part is capital income associated with the possession of an own house, the second part is all remaining capital income. As from 2001 the tax system has changed, as from then we define noncapital income as:

$$\text{net noncapital income} = L + T - \frac{L+T}{L+T+H}\tau_i - P + \text{allowances},$$

The difference with 1989-2000 is that $C$ is no longer in the definition. This is because the taxation on income does not include the income on capital (interests, dividend etc.) anymore. As from 2001 we define capital income as

$$\text{net capital income} = (H - \frac{H}{L+T+H}\tau_i) + (C - \tau_w),$$

also here the definition has changed because the taxation on income does not include capital income anymore.

## C Estimation method

The estimation of the model described in section 4 can be explained in two steps. In the first step the parameters of (1) are estimated. The second step investigates the autoregressive scheme.

To estimate the parameters $\beta$ of equation 1 we first compute

$$y_{it} - \bar{y}_i = \beta(x_{it} - \bar{x}_i) + (v_{it} - \bar{v}_i), \tag{7}$$

where $\bar{y}_i$ is the average household income of household $i$ across time. $\bar{x}_i$ and $\bar{v}_i$ are average values across time for each household $i$. Using (7), we can consistently estimate $\beta$ with OLS. Furthermore, we correct the standard errors of $\beta$ for the fact that the error terms $v_{it} - \bar{v}_i$ 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 $(\sum_{i=1}^{N}(T_i) - K)/(\sum_{i=1}^{N}(T_i - 1) - K)$ .

Averaging (1) across all observations gives

$$\bar{y} = \alpha + \beta\bar{x} + \bar{v}, \tag{8}$$

when we use the identifying assumption $\sum_{i=1}^{N}\mu_i = 0$. From this, we obtain $\hat{\alpha}$ by

$$\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x}, \tag{9}$$

We also average the data of individual households across time

$$\bar{y}_i = \alpha + \beta\bar{x}_i + \mu_i + \bar{v}_i, \tag{10}$$

from (10) we compute $\hat{\mu}_i = \bar{y}_i - \hat{\alpha} - \hat{\beta}\bar{x}_i$. Now $\hat{v}_{it}$ is computed by

$$\hat{v}_{it} = y_{it} - \hat{\alpha} - \hat{\beta}x_{it} - \hat{\mu}_i. \tag{11}$$

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

## D Estimation results transition models

This appendix describes the estimation results of the transition models with regard to marital status, children, and labour market positions explained in section 4.2 and 4.3.

### D.1 Marital status

Table 3 presents the estimation results of the transition models with regard to marital status. The explanatory variables in this estimations are age, age squared, cohort, and cohort squared. We assume period effects to be negligible compared to age and cohort effects. The first part of Table 3 presents the results for the transition married-divorced. The probability of a divorce decreases with age for both men and women and as expected younger cohorts divorce more often than older cohorts. The second part of the table is about the results of the transition divorced-married. The probability to remarry after a divorce decreases with age and is smaller for younger cohorts than for older cohorts. Only for women, the probability to remarry increases up to the cohort born in 1946. The estimations for the transition unmarried-married show that the probability of marriage mostly decreases with age and that younger cohorts have a higher probability to marry than older cohorts. This may look strange, as it is commonly known that younger cohorts marry less often than older cohorts. However, this sign can be explained by younger cohorts marrying later in life than the older cohorts. Therefore, the probability of marriage between age 36-90 is relatively high for younger cohorts. The fourth part of the table shows the estimation results with regard to the transition widowed-married. For widow(er)s the probability to remarry decreases with age and is higher for younger than for older cohorts.

### D.2 Children

The fifth part of Table 3 presents the estimation results with regard to new children being born. For this estimation all households in the years 1989-2006 are selected and we determine, given the characteristics in $t-1$, whether a new child has entered the household during the next year. The probability of a new child decreases with age, and is higher for younger cohorts. The probability of a child is higher in households with a couple and in households where already one child is present. So, if there is already one child present, there is a relatively large probability of a second child after age 36. On the other hand, when there are already two or more children in the household, the probability of an extra child after age 36 is relatively low. The last part of the table presents the estimation results with regard to the probability of children moving out of their parental household. For this estimation we select all children in the Dutch Income Panel 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 behaviour with regard to moving out as the children between 2006 and 2007.[17] As expected, the probability to move

[17]Here, children are defined as all persons younger than 30 who are at least 18 years younger than the key person of a household

out of the household increases with age. Furthermore, female children have a higher probability to move out of the household than males.

### D.3 Labour market position

For the transitions in labour market positions we estimate two models, one for couples and one for singles. Table 4 shows the estimation results for singles. The first half of the table is about the transition from work to occupational pension or 'other' (being no labour and no occupational pension, for example the receivers of just unemployment benefits or state pension). The probabilities to keep on working in the labour market increase for new generations, especially for men at higher ages and for women until age 75. Transition probabilities from work to 'other' decrease for younger generations of men and women. Divorced men and women have a relatively high probability to transit from work to 'other'. For women, the number of children influences the transitions from work to 'other' significantly positive. For men children have no significant effect. The second half of Table 4 deals with the transition 'other' to work or occupational pension. The transition probability to work is significantly higher for divorced persons.

For couples we estimate a logit model with nine outcomes. These nine outcomes are listed in Table 5, together with their relative frequencies. The combination where both members of the couple work has increased between 1989 and 2007, from 23% to 40%. On the other hand, the combination where only the man works has decreased (column three). Over the years, the percentage of couples involved in occupational pensions has increased. Within all couples where the man receives an occupational pension, couples where the woman also receives an occupational pension become more established.

As the labour market positions of the members of a couple are interrelated we treat the trivariate outcomes of the two persons of a couple as 9 univariate outcomes, and model the transitions between these nine states. The explanatory variables are the age of the man and woman part of the couple and the cohort to which the man belongs (in combination with the age of both man and woman, the cohort of the woman is automatically known). Also here we add an interaction of age and cohort to allow for age patterns to be different for different cohorts. [18]

Finally, Table 6 shows the gender-specific estimation results of the multinomial logit model to determine the inital labour market status of new household members and children who enter adulthood. All household members in the data are used. For men and women the probability of labour increases until about age 40 and decreases afterwards. The probability of receiving an occupational pension, on the other hand, increases with age.

[18]Because of the amount of space, the estimation results with regard to the transitions between these nine states are available on request.

Younger cohorts have a higher probability for a labour or occupational pensions status then older generations.

## E Theil decomposition

Using a Theil decomposition, overall income inequality can be related to mutually exclusive population subgroups. 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. Thus, 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) the households with occupational pension income and (2) the households without occupational pension income. The Theil index

$$T = \frac{1}{N} \sum_{i=1}^{N} \frac{y_i}{\bar{y}} \log(\frac{y_i}{\bar{y}}) \tag{12}$$

where $N$ is the number of observations, can be rewritten as

$$T = (s_1 T_1 + s_2 T_2) + (s_1 \log(\frac{\bar{y_1}}{\bar{y}}) + s_2 \log(\frac{\bar{y_2}}{\bar{y}})) \tag{13}$$

where the first term is within group inequality and the second term is between group inequality. $T_k$ is the Theil index for subgroup $k$, and $s_k$ is the share of total income received by subgroup $k$. $\bar{y_1}$ is the average income of households with occupational pension income and $\bar{y_2}$ is the average income of households without occupational pension income.

Table 3: Transition models marital status and children

| | Men | | Women | |
|---|---|---|---|---|
| | Coef | SE | Coef | SE |
| Married → Divorced | | | | |
| age/10 | 1.461 | 0.0414 | 1.611 | 0.0482 |
| $(age/10)^2$ | -0.190 | 0.0043 | -0.224 | 0.0052 |
| (year of birth-1900)/10 | 0.179 | 0.0404 | -0.097 | 0.0477 |
| $((year of birth-1900)/10)^2$ | 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 |
| $(age/10)^2$ | -0.032 | 0.0057 | -0.098 | 0.0074 |
| (year of birth-1900)/10 | -0.068 | 0.0539 | 0.290 | 0.0708 |
| $((year of birth-1900)/10)^2$ | -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 |
| $(age/10)^2$ | 0.203 | 0.0086 | 0.124 | 0.0123 |
| (year of birth-1900)/10 | 1.654 | 0.0919 | 1.927 | 0.1305 |
| $((year of birth-1900)/10)^2$ | -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 |
| $(age/10)^2$ | -0.175 | 0.0114 | -0.188 | 0.0130 |
| (year of birth-1900)/10 | 0.204 | 0.0919 | 0.419 | 0.1101 |
| $((year of birth-1900)/10)^2$ | -0.026 | 0.0106 | -0.018 | 0.0118 |
| constant | -5.406 | 0.3319 | -7.481 | 0.3376 |
| pseudo $R^2$ | 0.092 | | 0.143 | |
| New children being born | | | | |
| age/10 | -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 | | | |
| Children not moving out of the household (boys and girls together) | | | | |
| age | -0.008 | 0.0147 | | |
| $age^2$ | -0.005 | 0.0004 | | |
| gender | 0.474 | 0.0445 | | |
| constant | 4.072 | 0.1198 | | |
| pseudo $R^2$ | 0.165 | | | |
| N | 38906 | | | |

Table 4: Transition models labour market status for singles

| **Men** | Labour → Occup. pension | | Labour → Other | |
|---|---|---|---|---|
| | Coef | SE | Coef | SE |
| age/10 | 0.234 | 0.6802 | -1.194 | 0.2637 |
| $(age/10)^2$ | 0.091 | 0.0457 | 0.114 | 0.0218 |
| (year of birth-1900)/10 | -0.139 | 0.2471 | -0.314 | 0.0634 |
| interaction age and cohort | -0.020 | 0.0457 | -0.009 | 0.0206 |
| divorced | 0.240 | 0.0802 | 0.303 | 0.0538 |
| widow(er) | 0.345 | 0.1096 | -0.077 | 0.1227 |
| constant | -6.315 | 2.3818 | 1.601 | 0.6178 |
| pseudo $R^2$ | 0.063 | | | |
| N | 127759 | | | |
| **Women** | Labour → Occup. pension | | Labour → Other | |
| | Coef | SE | Coef | SE |
| age/10 | 0.121 | 0.7283 | -1.001 | 0.2838 |
| $(age/10)^2$ | 0.020 | 0.0476 | 0.090 | 0.0227 |
| (year of birth-1900)/10 | -0.812 | 0.2743 | -0.341 | 0.0733 |
| interaction age and cohort | 0.116 | 0.0488 | -0.041 | 0.0218 |
| divorced | -0.414 | 0.0800 | 0.454 | 0.0592 |
| widow(er) | 0.797 | 0.0805 | 0.048 | 0.1065 |
| # children | -0.099 | 0.0659 | 0.137 | 0.0202 |
| constant | -3.256 | 2.6181 | 1.954 | 0.7104 |
| pseudo $R^2$ | 0.104 | | | |
| N | 83337 | | | |
| **Men** | Other → Labour | | Other → Occup. pension | |
| | Coef | SE | Coef | SE |
| age/10 | -0.760 | 0.2741 | 2.067 | 0.8009 |
| $(age/10)^2$ | 0.017 | 0.0227 | -0.133 | 0.0513 |
| (year of birth-1900)/10 | 0.139 | 0.0644 | 0.205 | 0.3114 |
| interaction age and cohort | -0.045 | 0.0219 | -0.046 | 0.0524 |
| divorced | 0.502 | 0.0590 | 0.318 | 0.0907 |
| widow(er) | 0.219 | 0.1410 | 0.642 | 0.1074 |
| constant | 0.803 | 0.6192 | -10.647 | 2.9932 |
| pseudo $R^2$ | 0.183 | | | |
| N | 40821 | | | |
| **Women** | Other → Labour | | Other → Occup. pension | |
| | Coef | SE | Coef | SE |
| age/10 | -0.486 | 0.2744 | 0.680 | 0.6978 |
| $(age/10)^2$ | -0.017 | 0.0222 | -0.090 | 0.0426 |
| (year of birth-1900)/10 | 0.254 | 0.0684 | -0.882 | 0.2963 |
| interaction age and cohort | -0.065 | 0.0218 | 0.139 | 0.0456 |
| divorced | 0.582 | 0.0579 | -0.222 | 0.0795 |
| widow(er) | 0.441 | 0.1036 | 0.960 | 0.0718 |
| # children | -0.014 | 0.0169 | -0.113 | 0.0546 |
| constant | -0.164 | 0.6538 | -3.676 | 2.7678 |
| pseudo $R^2$ | 0.245 | | | |
| N | 49339 | | | |

This table shows the estimation results of the transition models of labour 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 5: Relative frequencies of the labour market positions of couples

| Year | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|
| 1989 | 22.8 | 1.1 | 36.0 | 1.6 | 1.5 | 14.7 | 4.1 | 0.8 | 17.4 |
| 1990 | 24.2 | 1.2 | 34.1 | 1.9 | 1.6 | 15.7 | 4.1 | 0.8 | 16.5 |
| 1991 | 25.6 | 1.2 | 32.2 | 1.8 | 1.9 | 16.2 | 4.2 | 0.8 | 16.1 |
| 1992 | 26.3 | 1.2 | 30.5 | 1.8 | 2.0 | 17.2 | 4.5 | 0.8 | 15.7 |
| 1993 | 27.7 | 1.2 | 29.4 | 2.0 | 2.2 | 17.3 | 4.6 | 0.8 | 15.0 |
| 1994 | 28.1 | 1.2 | 28.3 | 2.1 | 2.4 | 17.6 | 4.8 | 0.8 | 14.7 |
| 1995 | 29.3 | 1.3 | 27.6 | 2.1 | 2.6 | 17.5 | 4.9 | 0.8 | 13.9 |
| 1996 | 30.4 | 1.4 | 26.7 | 2.1 | 2.9 | 17.6 | 4.9 | 0.8 | 13.2 |
| 1997 | 31.6 | 1.4 | 25.9 | 2.3 | 3.2 | 17.8 | 4.9 | 0.7 | 12.3 |
| 1998 | 33.4 | 1.5 | 24.8 | 3.4 | 2.9 | 16.9 | 4.6 | 0.7 | 11.8 |
| 1999 | 35.6 | 1.4 | 23.8 | 3.5 | 2.7 | 17.0 | 4.5 | 0.7 | 10.8 |
| 2000 | 36.9 | 1.7 | 22.0 | 3.4 | 3.4 | 17.3 | 4.5 | 0.7 | 10.1 |
| 2001 | 36.0 | 1.5 | 21.7 | 2.6 | 3.9 | 19.4 | 4.3 | 0.6 | 9.9 |
| 2002 | 38.4 | 1.8 | 19.9 | 2.7 | 4.3 | 18.0 | 4.9 | 0.7 | 9.2 |
| 2003 | 38.5 | 1.9 | 18.9 | 3.0 | 4.7 | 18.0 | 5.3 | 0.8 | 9.0 |
| 2004 | 38.8 | 2.0 | 18.2 | 3.3 | 5.1 | 18.0 | 5.2 | 0.8 | 8.6 |
| 2005 | 38.8 | 2.0 | 17.5 | 3.5 | 5.4 | 18.1 | 5.5 | 0.8 | 8.4 |
| 2006 | 39.8 | 2.0 | 16.4 | 3.8 | 6.0 | 18.1 | 5.5 | 0.8 | 7.7 |
| 2007 | 40.3 | 1.9 | 15.5 | 4.1 | 6.5 | 18.2 | 5.4 | 0.8 | 7.2 |

Relative frequencies of the labour market status for 1989-2007. We distinguish three states: (1) labour, (2) occupational pension, (3) other (none of these two). In a couple we thus have 9 (=3*3) states. 1: man works, woman works, 2: man works, woman occ. pension, 3: man works, woman other, 4: man occ. pension, woman works, 5: man occ. pension, woman occ. pension, 6: man occ. pension, woman other, 7: man other, woman works, 8: man other, woman occ. pension 9: man other, woman other.

Table 6: Multinomial logit for the initial labour market status for new household members and children entering adulthood

| | Men | | Women | |
|---|---|---|---|---|
| | Coef | SE | Coef | SE |
| *Labour* | | | | |
| age/10 | 8.778 | 0.0404 | 6.150 | 0.0371 |
| $(age/10)^2$ | -1.768 | 0.0105 | -1.134 | 0.0095 |
| $(age/10)^3$ | 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 |
| $(age/10)^2$ | 1.548 | 0.0221 | 0.248 | 0.0251 |
| $(age/10)^3$ | -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 | |

Estimation results to predict initial labour market status of new household members (because of marriage) and children entering adulthood. We distinguish three labour market states: labour, occupational pension, and other (no labour and no occupational pension income). ‘Other’ is the reference category in this estimation.

# F Tables and figures

Table 7: Data selection

| | |
|---|---|
| Raw sample | 1,835,819 |
| | |
| *Observations left over after removal (sequentially)* | |
| Household income missing | 1,819,048 |
| Age of a household member missing | 1,819,007 |
| Negative or zero household income | 1,807,963 |
| Households with 9 or more household members | 1,802,405 |
| Key persons member of multiple couple household | 1,793,807 |
| Key person is a child or a student | 1,290,226 |
| Select key persons between 36-90 | 958,188 |
| Select key person born between 1917-1970 | 911,079 |
| Bottom or top 0.1% of income distribution (by year) | 909,257 |
| Minus the year 2000 after revision[a] | 861,336 |

The number of key persons and reason of removal from the sample. Key persons are randomly drawn from the Dutch population and are followed over time. We have information about all household members of the key persons.

[a] In the estimations we use the year 2000 before revision, instead of the year 2000 after revision.

Table 8: Descriptives equivalised household income, age key person 50-64

| Year | Mean | p10 | p50 | p90 | $\frac{p90}{p10}$ | $\frac{p90}{p50}$ | $\frac{p50}{p10}$ | Gini |
|---|---|---|---|---|---|---|---|---|
| 1989 | 20114 | 11310 | 18346 | 30705 | 2.71 | 1.67 | 1.62 | 0.228 |
| 1990 | 21187 | 11599 | 19096 | 32811 | 2.83 | 1.72 | 1.65 | 0.241 |
| 1991 | 21220 | 11464 | 19139 | 32566 | 2.84 | 1.70 | 1.67 | 0.243 |
| 1992 | 21183 | 11473 | 19242 | 32495 | 2.83 | 1.69 | 1.68 | 0.241 |
| 1993 | 21329 | 11530 | 19360 | 32931 | 2.86 | 1.70 | 1.68 | 0.241 |
| 1994 | 21241 | 11200 | 19210 | 33107 | 2.96 | 1.72 | 1.72 | 0.247 |
| 1995 | 21718 | 11320 | 19490 | 34049 | 3.01 | 1.75 | 1.72 | 0.250 |
| 1996 | 21971 | 11477 | 19727 | 34343 | 2.99 | 1.74 | 1.72 | 0.251 |
| 1997 | 22073 | 11530 | 19943 | 34418 | 2.99 | 1.73 | 1.73 | 0.248 |
| 1998 | 22747 | 12025 | 20534 | 35206 | 2.93 | 1.71 | 1.71 | 0.246 |
| 1999 | 23034 | 11985 | 20747 | 35923 | 3.00 | 1.73 | 1.73 | 0.253 |
| 2000 | 23596 | 12297 | 21190 | 36589 | 2.98 | 1.73 | 1.72 | 0.253 |
| 2000 | 23506 | 12428 | 21128 | 35947 | 2.89 | 1.70 | 1.70 | 0.248 |
| 2001 | 24203 | 12838 | 21786 | 37468 | 2.92 | 1.72 | 1.70 | 0.247 |
| 2002 | 24407 | 13024 | 22077 | 37580 | 2.89 | 1.70 | 1.70 | 0.244 |
| 2003 | 24128 | 12930 | 21911 | 37330 | 2.89 | 1.70 | 1.69 | 0.243 |
| 2004 | 24463 | 13124 | 22035 | 37641 | 2.87 | 1.71 | 1.68 | 0.245 |
| 2005 | 24589 | 13102 | 21994 | 38118 | 2.91 | 1.73 | 1.68 | 0.247 |
| 2006 | 23629 | 12598 | 20859 | 36872 | 2.93 | 1.77 | 1.66 | 0.254 |
| 2007 | 24351 | 12814 | 21528 | 38257 | 2.99 | 1.78 | 1.68 | 0.258 |

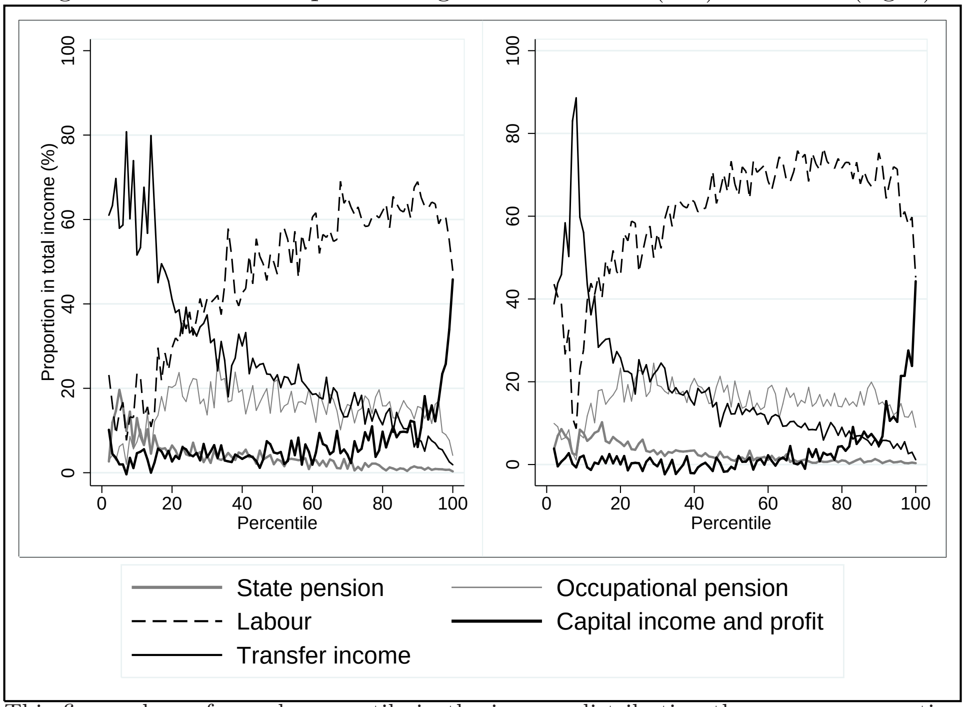
Source: IPO, own computations. In this paper income is always inflated/deflated to 2005 euro's. The year 2000 is presented two times, first for the data before revision and secondly for the data after revision.

Table 9: Descriptives equivalised household income, age key person 65-90

| Year | Mean | p10 | p50 | p90 | $\frac{p90}{p10}$ | $\frac{p90}{p50}$ | $\frac{p50}{p10}$ | Gini |
|---|---|---|---|---|---|---|---|---|
| 1989 | 17031 | 10355 | 14699 | 26732 | 2.58 | 1.82 | 1.42 | 0.225 |
| 1990 | 17725 | 10416 | 14850 | 28459 | 2.73 | 1.92 | 1.43 | 0.242 |
| 1991 | 17738 | 10388 | 14890 | 28641 | 2.76 | 1.92 | 1.43 | 0.244 |
| 1992 | 17626 | 10542 | 14935 | 28176 | 2.67 | 1.89 | 1.42 | 0.236 |
| 1993 | 17489 | 10557 | 14867 | 27746 | 2.63 | 1.87 | 1.41 | 0.231 |
| 1994 | 17252 | 10481 | 14639 | 27183 | 2.59 | 1.86 | 1.40 | 0.231 |
| 1995 | 17278 | 10605 | 14659 | 27246 | 2.57 | 1.86 | 1.38 | 0.228 |
| 1996 | 17375 | 10665 | 14799 | 27300 | 2.56 | 1.84 | 1.39 | 0.228 |
| 1997 | 17461 | 10835 | 14795 | 27343 | 2.52 | 1.85 | 1.37 | 0.225 |
| 1998 | 17916 | 11275 | 15192 | 27758 | 2.46 | 1.83 | 1.35 | 0.221 |
| 1999 | 17936 | 11222 | 15196 | 27696 | 2.47 | 1.82 | 1.35 | 0.224 |
| 2000 | 18337 | 11393 | 15515 | 28406 | 2.49 | 1.83 | 1.36 | 0.228 |
| 2000 | 18541 | 11504 | 15708 | 28337 | 2.46 | 1.80 | 1.37 | 0.227 |
| 2001 | 18562 | 11702 | 15737 | 28252 | 2.41 | 1.80 | 1.34 | 0.224 |
| 2002 | 19044 | 11932 | 16125 | 29300 | 2.46 | 1.82 | 1.35 | 0.225 |
| 2003 | 19065 | 11956 | 16180 | 29193 | 2.44 | 1.80 | 1.35 | 0.225 |
| 2004 | 19189 | 12073 | 16366 | 29316 | 2.43 | 1.79 | 1.36 | 0.222 |
| 2005 | 19367 | 12014 | 16439 | 29717 | 2.47 | 1.81 | 1.37 | 0.228 |
| 2006 | 19575 | 12247 | 16771 | 29788 | 2.43 | 1.78 | 1.37 | 0.224 |
| 2007 | 20048 | 12406 | 17196 | 30592 | 2.47 | 1.78 | 1.39 | 0.227 |

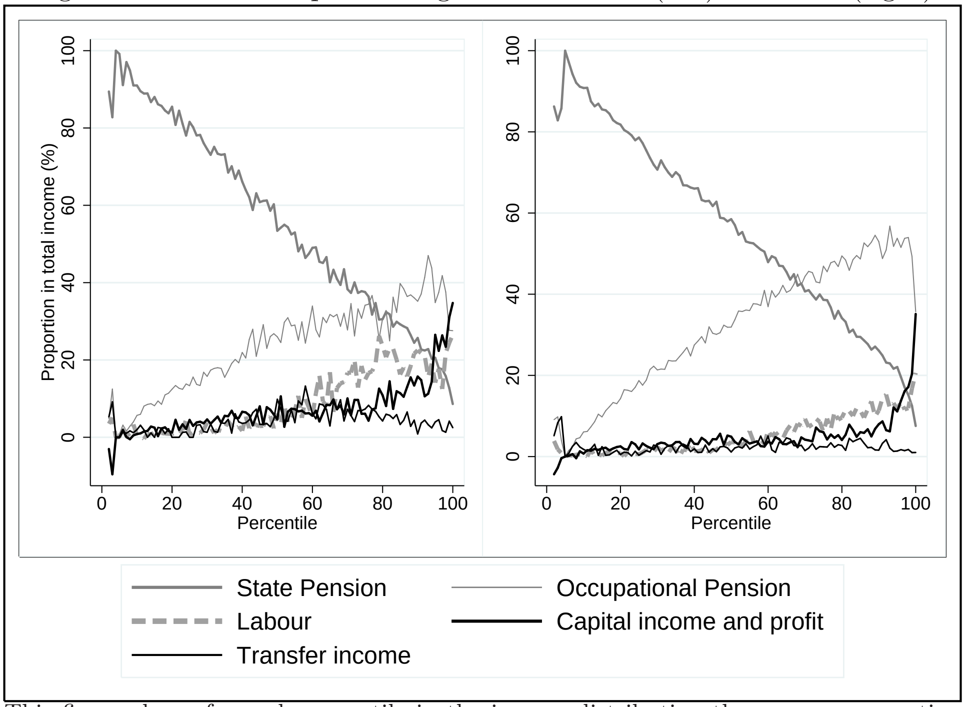
Source: IPO, own computations. In this paper income is always inflated/deflated to 2005 euro's. The year 2000 is presented two times, first for the data before revision and secondly for the data after revision.

Figure 6: Income composition age 50-64 in 1989 (left) and 2007 (right)




This figure shows for each percentile in the income distribution the average proportion of several income components. Transfer income includes welfare, disability benefits and unemployment benefits. Occupational pensions also contains early retirement income.

Figure 7: Income composition age 65-90 in 1989 (left) and 2007 (right)




This figure shows for each percentile in the income distribution the average proportion of several income components. Transfer income includes welfare, disability benefits and unemployment benefits. Occupational pensions also contains early retirement income.

Figure 8: Female participation in elderly couples where partner is 65 or older


In this figure we have selected all couples in which the man has reached age 65 and the woman is younger than 65. The figure shows the percentage of these couples in which the woman receives labour income.

Table 10: Transitions in marital status

| 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 |

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 age 36-90 and year of birth 1917-1970, own computations.

Table 11: Estimation results income equation

| | Coef 1 | SE 1 | Coef 2 | SE 2 | Coef 3 | SE 3 |
|---|---|---|---|---|---|---|
| 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 |
| 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 |

Estimation results continue on the next page.

Estimation results income equation, extended

| | **Coef 1** | **SE 1** | **Coef 2** | **SE 2** | **Coef 3** | **SE 3** |
|---|---|---|---|---|---|---|
| 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 |
| # 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 |
| # labour (man) | | | | | 0.120 | 0.0026 |
| # labour (woman) | | | | | 0.118 | 0.0018 |
| # occ. pension (man) | | | | | 0.058 | 0.0032 |
| # occ. pension (woman) | | | | | 0.099 | 0.0034 |
| $\rho_{0,1}$ | -0.160 | 0.0253 | -0.217 | 0.0259 | -0.227 | 0.0260 |
| $\rho_{1,1}$ | 0.162 | 0.0091 | 0.163 | 0.0093 | 0.162 | 0.0093 |
| $\rho_{2,1}$ | -0.010 | 0.0008 | -0.009 | 0.0008 | -0.009 | 0.0008 |
| $\rho_2$ | 0.066 | 0.0012 | 0.054 | 0.0012 | 0.055 | 0.0012 |
| $\alpha$ | 9.909 | | 9.746 | | 9.805 | |
| $\sigma_\mu$ | 0.370 | | 0.369 | | 0.342 | |
| $\sigma_\epsilon$ | 0.210 | | 0.205 | | 0.234 | |
| $R^2$ | 0.061 | | 0.134 | | 0.1645 | |
| N | 861336 | | 861336 | | 861336 | |

This table shows the estimation results of the income equation. Also year dummies are included in the estimation to account for period effects. 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.

Figure 9: Scheme of the cohorts used in the simulation to predict the income distribution of households with key person of age 50-90 in 2020

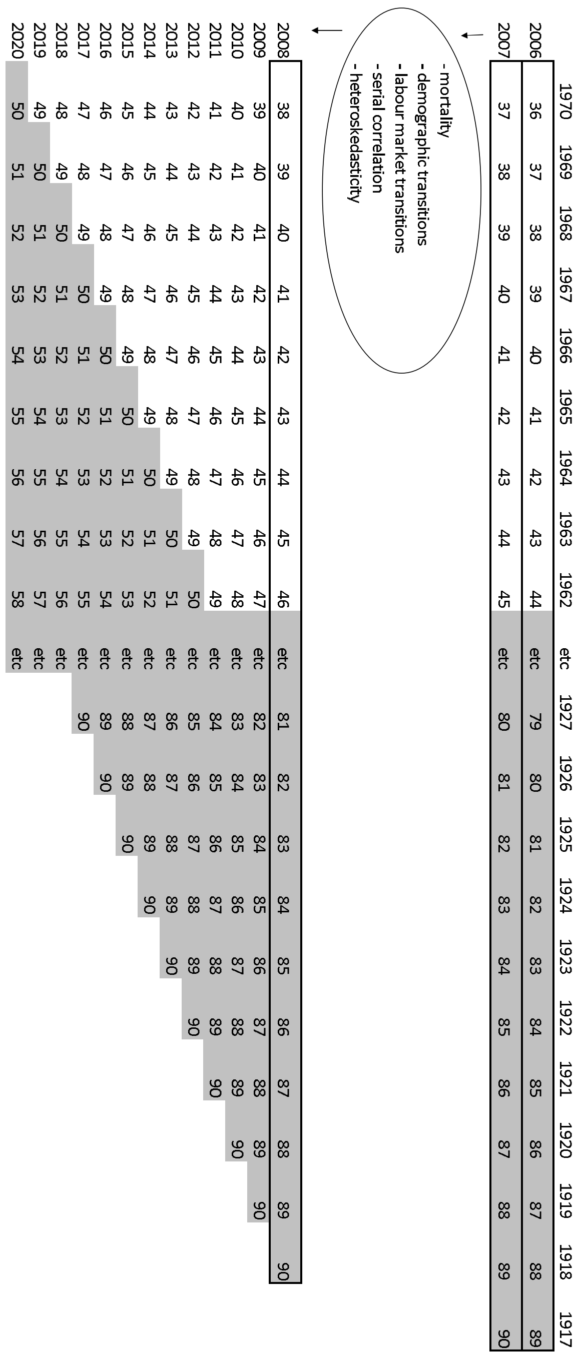


Table 12: Predicted marital status, age 50-64

| | **Men** | | | | **Women** | | | |
|---|---|---|---|---|---|---|---|---|
| Year | Marr | Unmarr | Wid | Div | Marr | Unmarr | Wid | Div |
| 2008 | 75.6 | 9.9 | 2.0 | 12.5 | 71.8 | 7.5 | 6.2 | 14.5 |
| 2009 | 74.7 | 10.5 | 2.0 | 12.8 | 70.9 | 7.9 | 6.1 | 15.1 |
| 2010 | 73.4 | 11.5 | 2.0 | 13.0 | 70.3 | 8.4 | 5.9 | 15.3 |
| 2011 | 72.2 | 12.1 | 2.0 | 13.7 | 69.5 | 9.0 | 5.7 | 15.7 |
| 2012 | 71.1 | 12.9 | 2.0 | 14.0 | 68.9 | 9.6 | 5.4 | 16.1 |
| 2013 | 70.1 | 13.6 | 1.9 | 14.4 | 68.0 | 10.1 | 5.5 | 16.4 |
| 2014 | 69.1 | 14.4 | 1.8 | 14.7 | 67.1 | 10.7 | 5.2 | 17.0 |
| 2015 | 67.7 | 15.4 | 1.8 | 15.1 | 66.3 | 11.3 | 5.0 | 17.4 |
| 2016 | 66.4 | 16.3 | 1.9 | 15.5 | 65.2 | 12.1 | 5.0 | 17.7 |
| 2017 | 65.4 | 17.0 | 1.9 | 15.8 | 64.4 | 12.7 | 4.9 | 18.0 |
| 2018 | 64.3 | 17.8 | 1.9 | 15.9 | 63.7 | 13.3 | 4.8 | 18.3 |
| 2019 | 63.1 | 18.8 | 1.7 | 16.4 | 62.8 | 14.2 | 4.6 | 18.4 |
| 2020 | 61.8 | 19.9 | 1.7 | 16.7 | 61.8 | 15.1 | 4.4 | 18.7 |

Table 13: Predicted marital status, age 65-90

| | **Men** | | | | **Women** | | | |
|---|---|---|---|---|---|---|---|---|
| Year | Marr | Unmarr | Wid | Div | Marr | Unmarr | Wid | Div |
| 2008 | 74.5 | 5.6 | 12.4 | 7.5 | 46.4 | 5.7 | 39.5 | 8.4 |
| 2009 | 73.8 | 5.7 | 12.6 | 7.9 | 46.9 | 5.5 | 39.0 | 8.6 |
| 2010 | 73.7 | 5.7 | 12.4 | 8.2 | 47.9 | 5.5 | 37.7 | 8.9 |
| 2011 | 73.4 | 5.8 | 12.2 | 8.7 | 48.2 | 5.5 | 36.9 | 9.4 |
| 2012 | 73.0 | 5.8 | 12.1 | 9.2 | 49.1 | 5.3 | 35.7 | 9.8 |
| 2013 | 72.4 | 5.9 | 12.1 | 9.6 | 49.8 | 5.4 | 34.7 | 10.2 |
| 2014 | 71.9 | 5.9 | 12.2 | 10.0 | 50.4 | 5.3 | 33.8 | 10.5 |
| 2015 | 71.7 | 6.0 | 12.0 | 10.2 | 50.4 | 5.3 | 33.3 | 11.0 |
| 2016 | 71.0 | 6.3 | 12.2 | 10.5 | 50.5 | 5.3 | 32.8 | 11.4 |
| 2017 | 70.9 | 6.4 | 12.1 | 10.6 | 50.6 | 5.5 | 32.2 | 11.7 |
| 2018 | 70.3 | 6.6 | 12.0 | 11.1 | 50.7 | 5.6 | 31.5 | 12.2 |
| 2019 | 69.1 | 7.1 | 12.4 | 11.4 | 50.5 | 5.8 | 31.1 | 12.7 |
| 2020 | 68.4 | 7.4 | 12.4 | 11.8 | 50.2 | 5.9 | 30.8 | 13.2 |

Table 14: Predicted labour market status age 50-64

| | | **Man** | | | **Woman** | |
|---|---|---|---|---|---|---|
| Year | Labour | Occupational pension | Other | Labour | Occupational pension | Other |
| 2008 | 62.6 | 19.6 | 17.8 | 46.3 | 15.0 | 38.6 |
| 2009 | 62.8 | 21.2 | 16.1 | 47.4 | 16.6 | 36.0 |
| 2010 | 62.3 | 23.0 | 14.7 | 48.4 | 17.9 | 33.7 |
| 2011 | 63.1 | 23.3 | 13.7 | 49.9 | 19.0 | 31.1 |
| 2012 | 63.3 | 23.6 | 13.1 | 51.9 | 19.8 | 28.2 |
| 2013 | 64.4 | 23.7 | 11.9 | 53.6 | 20.6 | 25.8 |
| 2014 | 64.5 | 24.2 | 11.3 | 55.1 | 21.4 | 23.5 |
| 2015 | 64.6 | 24.6 | 10.8 | 56.1 | 22.4 | 21.6 |
| 2016 | 65.0 | 25.0 | 10.1 | 56.7 | 23.5 | 19.8 |
| 2017 | 65.3 | 25.3 | 9.5 | 57.4 | 24.3 | 18.3 |
| 2018 | 65.8 | 25.0 | 9.2 | 58.1 | 25.1 | 16.7 |
| 2019 | 66.2 | 25.1 | 8.7 | 59.1 | 25.7 | 15.2 |
| 2020 | 66.3 | 25.4 | 8.4 | 59.5 | 26.7 | 13.9 |

In case a person receives both labour income and occupational pension income the labour market status is based on the highest income component.

Table 15: Predicted labour market status age 65-90

| | | **Man** | | | **Woman** | |
|---|---|---|---|---|---|---|
| Year | Labour | Occupational pension | Other | Labour | Occupational pension | Other |
| 2008 | 3.6 | 87.0 | 9.4 | 2.1 | 54.0 | 43.8 |
| 2009 | 3.2 | 87.5 | 9.3 | 2.0 | 54.8 | 43.1 |
| 2010 | 3.1 | 88.0 | 8.9 | 2.2 | 55.3 | 42.5 |
| 2011 | 3.6 | 88.0 | 8.4 | 2.6 | 56.1 | 41.3 |
| 2012 | 4.3 | 87.6 | 8.1 | 3.0 | 56.9 | 40.2 |
| 2013 | 4.2 | 88.1 | 7.7 | 3.2 | 58.5 | 38.3 |
| 2014 | 4.1 | 88.8 | 7.1 | 3.1 | 60.1 | 36.8 |
| 2015 | 4.4 | 88.9 | 6.6 | 2.9 | 62.0 | 35.1 |
| 2016 | 4.3 | 89.5 | 6.2 | 2.9 | 63.9 | 33.2 |
| 2017 | 4.1 | 89.8 | 6.1 | 3.0 | 65.4 | 31.6 |
| 2018 | 4.3 | 90.0 | 5.6 | 3.0 | 67.3 | 29.7 |
| 2019 | 4.1 | 90.6 | 5.3 | 3.2 | 68.6 | 28.2 |
| 2020 | 4.0 | 90.9 | 5.1 | 3.1 | 70.4 | 26.5 |

In case a person receives both labour income and occupational pension income the labour market status is based on the highest income component.

Table 16: Descriptives predicted equivalised household income, age 50-64

| Year | Mean | p10 | p50 | p90 | $\frac{p90}{p10}$ | $\frac{p90}{p50}$ | $\frac{p50}{p10}$ | Gini |
|---|---|---|---|---|---|---|---|---|
| *Specification 1* | | | | | | | | |
| 2008 | 24340 | 13167 | 21905 | 37813 | 2.87 | 1.73 | 1.66 | 0.244 |
| 2009 | 24418 | 13205 | 21996 | 37792 | 2.86 | 1.72 | 1.67 | 0.240 |
| 2010 | 24456 | 13261 | 22066 | 37849 | 2.85 | 1.72 | 1.66 | 0.240 |
| 2011 | 24471 | 13318 | 22110 | 37783 | 2.84 | 1.71 | 1.66 | 0.237 |
| 2012 | 24609 | 13437 | 22288 | 37814 | 2.81 | 1.70 | 1.66 | 0.238 |
| 2013 | 24815 | 13544 | 22400 | 38242 | 2.82 | 1.71 | 1.65 | 0.239 |
| 2014 | 24877 | 13511 | 22547 | 38347 | 2.84 | 1.70 | 1.67 | 0.240 |
| 2015 | 25146 | 13562 | 22600 | 38960 | 2.87 | 1.72 | 1.67 | 0.244 |
| 2016 | 25215 | 13619 | 22666 | 39207 | 2.88 | 1.73 | 1.66 | 0.242 |
| 2017 | 25362 | 13692 | 22787 | 39536 | 2.89 | 1.74 | 1.66 | 0.243 |
| 2018 | 25627 | 13760 | 22999 | 40067 | 2.91 | 1.74 | 1.67 | 0.245 |
| 2019 | 25940 | 13869 | 23163 | 40706 | 2.94 | 1.76 | 1.67 | 0.248 |
| 2020 | 26047 | 13893 | 23180 | 40909 | 2.94 | 1.76 | 1.67 | 0.251 |
| *Specification 2* | | | | | | | | |
| 2008 | 24659 | 13199 | 22205 | 38418 | 2.91 | 1.73 | 1.68 | 0.245 |
| 2009 | 24752 | 13276 | 22332 | 38160 | 2.87 | 1.71 | 1.68 | 0.242 |
| 2010 | 24950 | 13394 | 22558 | 38531 | 2.88 | 1.71 | 1.68 | 0.241 |
| 2011 | 25083 | 13601 | 22643 | 38561 | 2.84 | 1.70 | 1.66 | 0.240 |
| 2012 | 25153 | 13721 | 22857 | 38765 | 2.83 | 1.70 | 1.67 | 0.237 |
| 2013 | 25426 | 13688 | 22902 | 39381 | 2.88 | 1.72 | 1.67 | 0.241 |
| 2014 | 25564 | 13711 | 23081 | 39328 | 2.87 | 1.70 | 1.68 | 0.241 |
| 2015 | 25793 | 13828 | 23283 | 40008 | 2.89 | 1.72 | 1.68 | 0.242 |
| 2016 | 25890 | 13891 | 23238 | 40142 | 2.89 | 1.73 | 1.67 | 0.244 |
| 2017 | 26039 | 14051 | 23483 | 40459 | 2.88 | 1.72 | 1.67 | 0.242 |
| 2018 | 26125 | 13939 | 23528 | 40733 | 2.92 | 1.73 | 1.69 | 0.245 |
| 2019 | 26414 | 14026 | 23618 | 41364 | 2.95 | 1.75 | 1.68 | 0.248 |
| 2020 | 26516 | 14068 | 23688 | 41657 | 2.96 | 1.76 | 1.68 | 0.249 |
| *Specification 3* | | | | | | | | |
| 2008 | 24484 | 13282 | 22028 | 37983 | 2.86 | 1.72 | 1.66 | 0.242 |
| 2009 | 24665 | 13473 | 22285 | 38117 | 2.83 | 1.71 | 1.65 | 0.237 |
| 2010 | 24789 | 13573 | 22443 | 38283 | 2.82 | 1.71 | 1.65 | 0.237 |
| 2011 | 24877 | 13757 | 22594 | 38348 | 2.79 | 1.70 | 1.64 | 0.235 |
| 2012 | 25054 | 13787 | 22760 | 38469 | 2.79 | 1.69 | 1.65 | 0.235 |
| 2013 | 25309 | 13986 | 22839 | 38960 | 2.79 | 1.71 | 1.63 | 0.238 |
| 2014 | 25506 | 14015 | 22961 | 39209 | 2.80 | 1.71 | 1.64 | 0.239 |
| 2015 | 25592 | 14062 | 23047 | 39493 | 2.81 | 1.71 | 1.64 | 0.239 |
| 2016 | 25672 | 13936 | 23004 | 39560 | 2.84 | 1.72 | 1.65 | 0.243 |
| 2017 | 25719 | 13937 | 23000 | 39649 | 2.84 | 1.72 | 1.65 | 0.243 |
| 2018 | 25953 | 13880 | 23264 | 40443 | 2.91 | 1.74 | 1.68 | 0.247 |
| 2019 | 26059 | 13929 | 23400 | 40473 | 2.91 | 1.73 | 1.68 | 0.248 |
| 2020 | 26239 | 13946 | 23387 | 41488 | 2.97 | 1.77 | 1.68 | 0.251 |

In this paper income is always inflated/deflated to 2005 euro's.

Table 17: Descriptives predicted equivalised household income, age 65-90

| Year | Mean | p10 | p50 | p90 | $\frac{p90}{p10}$ | $\frac{p90}{p50}$ | $\frac{p50}{p10}$ | Gini |
|---|---|---|---|---|---|---|---|---|
| *Specification 1* | | | | | | | | |
| 2008 | 20327 | 12178 | 17842 | 31186 | 2.56 | 1.75 | 1.47 | 0.227 |
| 2009 | 20634 | 12025 | 18162 | 32018 | 2.66 | 1.76 | 1.51 | 0.230 |
| 2010 | 20821 | 11985 | 18491 | 32459 | 2.71 | 1.76 | 1.54 | 0.231 |
| 2011 | 21159 | 12096 | 18866 | 32871 | 2.72 | 1.74 | 1.56 | 0.233 |
| 2012 | 21398 | 12140 | 19108 | 33252 | 2.74 | 1.74 | 1.57 | 0.232 |
| 2013 | 21646 | 12276 | 19435 | 33462 | 2.73 | 1.72 | 1.58 | 0.233 |
| 2014 | 21803 | 12272 | 19615 | 33734 | 2.75 | 1.72 | 1.60 | 0.231 |
| 2015 | 21975 | 12271 | 19822 | 33814 | 2.76 | 1.71 | 1.62 | 0.231 |
| 2016 | 22184 | 12419 | 19964 | 34220 | 2.76 | 1.71 | 1.61 | 0.230 |
| 2017 | 22365 | 12561 | 20116 | 34307 | 2.73 | 1.71 | 1.60 | 0.230 |
| 2018 | 22491 | 12660 | 20361 | 34504 | 2.73 | 1.69 | 1.61 | 0.228 |
| 2019 | 22635 | 12770 | 20554 | 34385 | 2.69 | 1.67 | 1.61 | 0.226 |
| 2020 | 22808 | 12826 | 20714 | 34901 | 2.72 | 1.68 | 1.62 | 0.228 |
| *Specification 2* | | | | | | | | |
| 2008 | 20315 | 12230 | 17843 | 31198 | 2.55 | 1.75 | 1.46 | 0.227 |
| 2009 | 20687 | 12140 | 18211 | 32134 | 2.65 | 1.76 | 1.50 | 0.231 |
| 2010 | 20925 | 12108 | 18501 | 32606 | 2.69 | 1.76 | 1.53 | 0.233 |
| 2011 | 21389 | 12191 | 18875 | 33440 | 2.74 | 1.77 | 1.55 | 0.236 |
| 2012 | 21707 | 12260 | 19203 | 33790 | 2.76 | 1.76 | 1.57 | 0.236 |
| 2013 | 21909 | 12287 | 19475 | 34233 | 2.79 | 1.76 | 1.59 | 0.236 |
| 2014 | 22188 | 12478 | 19702 | 34338 | 2.75 | 1.74 | 1.58 | 0.235 |
| 2015 | 22408 | 12604 | 19996 | 34609 | 2.75 | 1.73 | 1.59 | 0.234 |
| 2016 | 22644 | 12723 | 20373 | 35022 | 2.75 | 1.72 | 1.60 | 0.233 |
| 2017 | 22788 | 12819 | 20503 | 35225 | 2.75 | 1.72 | 1.60 | 0.231 |
| 2018 | 22954 | 12909 | 20728 | 35391 | 2.74 | 1.71 | 1.61 | 0.229 |
| 2019 | 23255 | 13035 | 20937 | 35665 | 2.74 | 1.70 | 1.61 | 0.231 |
| 2020 | 23425 | 13121 | 21156 | 35859 | 2.73 | 1.69 | 1.61 | 0.229 |
| *Specification 3* | | | | | | | | |
| 2008 | 20267 | 12214 | 17805 | 31156 | 2.55 | 1.75 | 1.46 | 0.225 |
| 2009 | 20611 | 12147 | 18122 | 31940 | 2.63 | 1.76 | 1.49 | 0.230 |
| 2010 | 20875 | 12280 | 18443 | 32162 | 2.62 | 1.74 | 1.50 | 0.229 |
| 2011 | 21252 | 12377 | 18862 | 32715 | 2.64 | 1.73 | 1.52 | 0.229 |
| 2012 | 21508 | 12437 | 19188 | 33144 | 2.66 | 1.73 | 1.54 | 0.229 |
| 2013 | 21754 | 12522 | 19396 | 33332 | 2.66 | 1.72 | 1.55 | 0.229 |
| 2014 | 21951 | 12734 | 19612 | 33716 | 2.65 | 1.72 | 1.54 | 0.227 |
| 2015 | 22212 | 12829 | 19890 | 34059 | 2.65 | 1.71 | 1.55 | 0.227 |
| 2016 | 22301 | 12860 | 19943 | 34105 | 2.65 | 1.71 | 1.55 | 0.228 |
| 2017 | 22436 | 12836 | 20111 | 34332 | 2.67 | 1.71 | 1.57 | 0.230 |
| 2018 | 22589 | 12944 | 20331 | 34457 | 2.66 | 1.69 | 1.57 | 0.229 |
| 2019 | 22672 | 12949 | 20434 | 34463 | 2.66 | 1.69 | 1.58 | 0.229 |
| 2020 | 22874 | 12976 | 20632 | 34925 | 2.69 | 1.69 | 1.59 | 0.230 |

Figure 10: Indexed growth of equivalised household income for retirees (age 65-90). The first figure is based on model specification 1, the second on model specification 2. Dashed lines are predictions, solid lines are realizations corrected for period effects.