The future development of living standards of the retirees in Belgium. [:] an application of the static microsimulation model station

Dekkers, gijs

Centre for Social Policy (CSB)

1999

The future development of living standards of the retirees in Belgium.

An application of the STAtic microsimulaTION model STATION

Gijs J.M. Dekkers
Centre for Social Policy (CSB)
Antwerp University UFSIA

November, 1999

Abstract: This paper develops a dynamic microsimulation model with static ageing to assess the consequences of the assumptions and hypothesis of the Federal Planning Bureau on the prospective adequacy of pensions. A less technical and shorter version of this text was published as Gijs Dekkers, 2000, L'évolution du pouvoir d'achat des retraités: Une application du modèle de microsimulation STATION. in: Pestieau, P., L. Gevers, V. Ginsburgh, E. Schokkaert, B. Cantillon, *Réflexions sur l'avenir de nos Retraites*, Garant, Leuven/Apeldoorn (also available in Dutch).

JEL-codes: J14, D31, I32

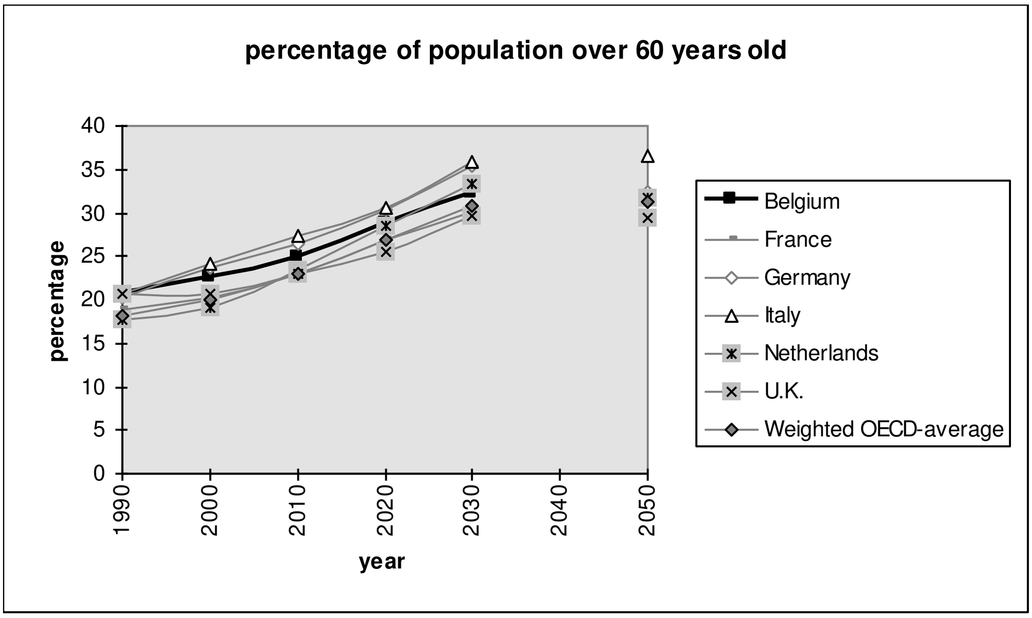
Keywords: microsimulation, ageing, social hypotheses, poverty, inequality

Chapter 1. Introduction and problem formulation.

1.1. Introduction.

As is the case in most western countries, the Belgian population is greying. As an illustration, consider the following figure, where the number of individuals over 60 years old is expressed as a percentage of the population (source: World Bank, 1994, P. 349).

Figure 1: Percentage of population over 60 years old.



The percentage of retirees is relatively high in Belgium in 1990, but is growing somewhat slower than in the other countries, so that it ends up somewhere in the middle in 2050. So, the rate of ageing may be moderate if compared to countries such as the Netherlands and Italy, but it remains significant in itself. As could be expected, the consequences of such important population shifts have been explored thoroughly, not only its macroeconomic consequences, but (and probably even primarily) its consequences for the financial sustainability of public pension systems and health care services. (for instance, see World Bank, 1994, Creedy (ed.) 1995 and Bos (ed), 1993, Jackson, 1992, Jackson, 1998, Lesthaeghe, Meeusen, Vandewalle, 1998, Quinn, 1997). As far as pensions are concerned, various empirical models have been developed with the aim of giving policy makers insight to the process of ageing and its consequences for the financial sustainability of pensions. (For instance, see, Huijzer and van Loo, 1986, Bolhuis and Vossers, 1986, Jansweijer, 1996 , van

Heerwaarden, Eikelboon, den Heijer, 1996, for the Netherlands, Bengtsson and Kruse, 1993 for Sweden, De Callatay and Turtelboom, 1996, and (in a broader context) d'Alcantara and Wardenier, 1997, Aaron, 1977, Kingsland, 1982).

The model which has recently been developed at the Belgian Federal Planning Bureau (Festjens, 1997), referred to as PENSION, fits in this rich tradition. As is the cast in most of the above models, it disentangles several subgroups within the ageing population and shows what will happen to the contribution rates if the average pension benefits are kept constant, or vice versa. It is a model of the 'flow or vintage type', where annual flows of 7 types of retirees (Festjens, idem, p. 6) are taken as a basis for a 'mechanical' calculation of the pension benefits (Festjens, idem, p. 5). The author refer to the model PENSION as a submodule of the general model MALTESE (Festjens, idem, p.7), thereby emphasising the tight links between PENSION and the other models developed at the Bureau. An important characteristic of the model is that it does not extrapolate historical information, since this information can not be a reference for future generations of retirees (Festjens, idem, p.4 paragraph B). To a certain extent, this is certainly true: specific circumstances (such as the second World-War or the economic heydays of the sixties) cause the socio-economic profile of generations to be different from other generations. However, it seems the most efficient to use all information available today, including historical information. Moreover, the model uses simulation results from macroeconomic models (HERMES and MALTESE) and demographic projections as input factors. These models combine historical information with scientific knowledge about identities and causal relations. Consequently, historical information does enter the pension models, though this link is indirectly. Lastly, and this is specific to the questions the model to be presented in this study tries to answer, the only information we have on the future distribution on income is the current distribution of income. Up to today, theory does not provide us with specific unambiguous empirical identities or causal relations which we can use to simulate the income distribution in a future point in time, without using the current distribution of income as a point of departure. We therefore have no choice but to use the historical information on the distribution of income.

The simulation results of these kinds of vintage-like models as PENSION is, usually take the form of time-series, showing the simulated future development of, say, contribution rates and pension-benefits. This allows policy makers to see whether or not ageing will become a (financing) problem in the future, and give them an idea on what actions they could take in order to preserve financial sustainability. Useful as these models are, they fail to show the redistributive impact of the pension system and the effect of ageing on this redistribution (and related variables, such as poverty). In other words, these models show what the policy makers can do to keep the pension system payable, if would they decide that action is required. But they do not show what the effect of these potential measures on the distribution of pension-benefits or the poverty rates among the retirees will be. It is however clear that when we consider the future welfare of the retirees, it is not enough only to look at the development of the average pension benefit. Other information, such as the distribution of (pension)income around the mean, poverty rates and such should be taken into account[1].

To overcome this lack of information concerning income distribution, microsimulation models have been developed. What microsimulation models exactly are will be explained in depth in the next paragraph, but let us just highlight the basic difference with the models we just came to mention. This difference is that the point of departure of these models is the population (or subgroups within the population) as a whole, where the point of departure of microsimulation models is the individual itself.

In this study, the effect of ageing on pension income inequality, poverty and welfare of the retirees will be considered. It this respect, the model which forms the basis of this study, the microsimulation model STATION, can be seen as being complementary to the PENSION-model of the Belgian Federal Planning Bureau, since it uses some of the assumptions of this model to show the consequences on the poverty rates, income distribution and future development of welfare of the retirees. So, this study concentrates on what effect potential policy measures (or the not-taking of these policy measures) will have on the future income distribution, poverty and welfare development of the retirees. The 'income-side' of the pension system is under consideration only to the extent that it is strictly necessary, so for questions as 'will

[1] Barry (1990, p.2) describes the link between welfare and (redistributional) justice as '*inextricably*'. Note that using income as a notion of household welfare is subject to criticism (see Slesnick, 1998), for instance because this assumes constant marginal utility of money and because price-changes are ignored.

the pension system remain payable in the future', the reader is referred to other literature, notably Festjens, 1997, and de Callatay et. al, 1996.

The above discussion can be summarised in the following problem definition: what is the effect of ageing on the development of pension-benefits, the inequality of pension-benefits and poverty and how will these circumstances change as a result of some assumptions and policy measures used in the PENSION-model by the Federal Planning Bureau?

This study starts by a brief discussion of the Belgian pension system. Next, we will turn our attention to microsimulation models: what is microsimulation? What kind of techniques are associated with microsimulation and what are their advantages and disadvantages? As a third step, the static microsimulation model STATION will be explained in depth and the fourth step will then be the presentation and discussion of the simulation results. Finally, conclusions will be drawn.

Before proceeding, a last introductory remark must be made. In order to facilitate co-operation with researchers from other universities, an internet-homepage has been made. The address of this STATION-homepage is *http://www.ufsia.ac.be/~gdekkers/station/index.htm*. This homepage offers those who know the required passwords the possibility to download the simulation results (i.e. the weighting variables), which can then be used in other empirical studies. Of course, a more extensive technical description of how to use the simulation results, is provided as well.

1.2. A birds-eye view on the Belgian pension system.

The Belgian pension system consists of three layers of which the first one is the most important. The first layer consists of three separate state-wide pension systems for employees, civil servants and self-employed. The benefits paid out to retirees in a certain year are financed by contributions of the working generations (a system widely known as a Pay-As-You-go System or PAYG), though the government contributes heavily as well[2]. For all three systems, the pension benefit is equal to 60 or

[2] For instance, in 1985, the (expected) contributions of the state covered 19.8% of total pension benefits paid out to former employees. For the self-employed and the civil-servants, these percentages were 32.3% and 71.2%, respectively (de Cock, 1984, page 56).

75%[3] of a certain wage base, times the relative length of the career (expressed in 45th for males and 40th for females, though this latter figure is now gradually being adjusted). The wage-base is either the career-long average wage (or profit) for employees and self-employed, or the average wage of only the last five years of the career for civil servants. Given a certain minimal length of career, the system provides a minimum pension benefit.

These separate pension schemes for employees, self-employed and civil servants are supplemented by a system of 'guaranteed income for retirees', a welfare-scheme providing those retirees who never had a career or a career of insufficient length with a minimal and means-tested pension benefit.

The second layer of the Belgian pension system consists of semi-collective additional pension schemes, organized on the firm level by pension funds. Though this second layer is rapidly gaining importance, the relative number of pension-receiving households is still rather limited (see Dekkers, 1998, table 2 and Neyt, 1993, p. 362). The third level consists of individual pension schemes and life-insurances. These last two pension benefits are voluntary, based on capital funding instead of Pay-As-You-Go and are complementary to the nation-wide first-layer pension system. The microsimulation model STATION, which is the *raison d'être* of this text, concentrates on the first layer of the Belgian pension system.

1.3. Demographic trends in Belgium

Next, let us glance at the demographical situation of Belgium for a moment. This paragraph, which draws heavily on chapters 2 and 3 of the book by Lesthaege, Meeusen and Vandewalle (idem, 1998), will briefly discuss both past and expected-future demographic trends in Belgium.

The recent 'demographic history' of Belgium is characterized by two important demographic transitions, of which the second is the most relevant in this context. The first demographic transition took place in the nineteenth century and started with a decrease of the mortality rate (as a result of medical improvements and increasing knowledge on hygiene). This was followed by a decrease of the fertility

[3] The latter is the family pension benefit, of which only one of the marital partners is eligible, i.e. given that the other partner refrains from his or her individual pension claim.

rate, resulting from urbanisation and secularisation. Moreover, the average age of marriage started to decrease, and the relative number of celibate individuals rose.

The second demographic roughly started in the second half of the fifties. This transition was triggered by changing social values, among other things on the role of the family in society and the emancipation of women. During the first half of the sixties, and probably partially resulting from the fact that economic growth rates were very high, the average age of marriage decreased even further, and so did the average number of years between marriage and having the first child. Consequently, fertility reached a maximum, resulting in the so-called 'baby-boom'. However, mainly as a result from the introduction of the contraceptive pill, combined with increasing economic independence of women (and therefore higher 'opportunity costs' of maternity), the fertility rate again decreased rapidly. This effect was strengthened by a delay of marriage and parentage. As a result, a 'baby-bust' started in the second half of the seventies and continued during the eighties. The fertility rate was 2.25 in 1970, which is above replacement-rate, but then decreased to 1.69 in 1980 and 1.55 in 1996 and 1996 (idem, 1998, table 2.6, p. 40).

Demographic behavioural changes like the ones described hitherto change demographic structures and the effects of this are of a typical long-term nature. These changes therefore form a basis for projections of the future. Additionally, assumptions on the future course of some key variables make it possible to distinguish simulation variants. In this text, only one variant will be discussed, since it forms the basis of the simulations of the Federal Planning Bureau and, consequently, our own simulations[4]. The key assumptions underlying the projections of scenario A of the National Institute of Statistics and the Federal Planning Bureau are, first of all, that the life expectancy at birth (which is now 80 years for men and women taken together) will increase to 82.1 for males and 88.1 for females. Secondly, the migration balance (which has a positive balance of 10,000 immigrants per year in 1995) will decrease to 3,000 in 2050. Thirdly, fertility increases rapidly from 1,55 in 1995 to 1.75 in 2010 and remains stable thereafter.

---

[4] This is variant A. This variant, together with the other variants, is discussed in Lesthaeghe et. al., 1998, p. 62 and further.

It must however be noted that Lesthaeghe et. al. think that the assumed future recovery of the fertility rate is too high, and that the decrease of the proportion of individuals younger than 20 is therefore underestimated.

The described historical developments, together with the assumptions, result in some main demographic trends. First of all, the rate of ageing will be quite strong between 2010 and 2030. It will not only be caused by an increase of the proportion of retirees (‘greying’), but by a decrease of the proportion of young individuals as well. Thirdly, immigration can prevent the population to decrease from a certain point on, but its negative effect on ageing will be small, if any.

Of course, the changes in demographic behaviour as mentioned above, have their consequences for the household structure in Belgium. Based on projections by Boulanger et. al. (idem, up to 2011), the following trends for four broad age-categories can be mentioned. As far as children and young individuals up to 20 years of age are considered, the most important trend is that the proportion of children living in households where there is only one parent, or where parents cohabit, will increase. Young adults (between 20 and 34) will tend to remain living with their parents more often. Moreover, the proportion of married individuals, especially with children, in this age-category will decrease. From the beginning of the eighties on, the proportion of young adults living alone has been increasing. This trend will persist, though at a lower speed as more and more individuals will cohabit.

For older adults (say between 40 and 65), the trend that the proportion of married parents decreases, emerges as well. This effect is however less strong for individuals of 55 and older, since more young adults postpone forming their own household and remain living with their parents. As a result of an increasing probability of divorce, the proportion of older adults who live alone, increases.

Lastly, the trends for the retirees must be described. Two main trends emerge: first of all, the life expectancy of couples increases strongly, so the average age of losing ones partner increases as well. Secondly, the proportion of retirees living with relatives will decrease, as will be the case with the proportion of retirees living in institutions.

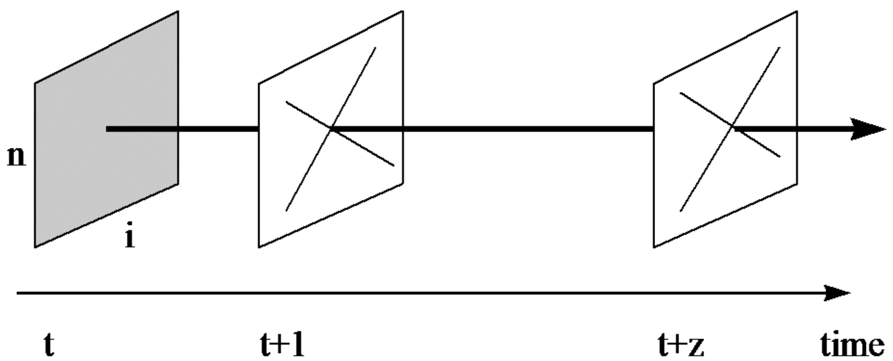
Here ends the description of the demographic situation of Belgium, now and in the future. Next, the question what microsimulation is and what can be done with it, will be answered. The following step will then be the presentation and description of the microsimulation model STATION and its simulation results.

1.4. What is microsimulation?

Socio-economic models can be subdivided according to the level on which they apply. First of all, there are macroeconomic models. These models simulate entire countries or even groups of countries. Secondly, there are meso-economic models which concentrate on the simulation of one or more branches of industry within a country. The third category of socio-economic models has emerged the most recently and take (groups of) individuals as the point of departure. The models in this category are called microsimulation models and aim at evaluating the effect of various economic- and social changes on the distribution of certain characteristics for different groups of individuals. Mostly , the goal of microsimulation models is to analyse the changes in the poverty rates and the distribution of income over groups in the population, resulting from external changes, such as demographic changes, economic development and policy changes.

The way which microsimulation models work can best be explained by rephrasing it to a problem common in econometrics and sociometrics, namely that of missing data analysis. Suppose we have a cross-sectional dataset at time t, consisting of n variables describing i individuals. Suppose furthermore a dataset of n variables and j individuals at the future time point $t+z$, $z>0$, which can be considered as consisting completely of missings, as shown in figure 2:

Figure 2: future data is missing...



Now microsimulation models basically are tools to fill in the missing datasets at the future time point $t+1$ up to $t+z$. Standard textbook econometrics learns us that missings can be filled in by two general methods: cold-deck imputation and hot-deck imputation (for an introduction to these techniques, see Kalton, 1983). These two methods form the basis of the division of microsimulation models in static[5] - and dynamic microsimulation models. Both types will be discussed briefly below.

1.4.1. Dynamic microsimulation.

Dynamic microsimulation basically fills in the missing datasets by using hot-deck imputation. Taking the cross-sectional dataset in time t as the point of departure, every individual in this dataset faces certain probabilities of a change in each of the n variables which describe him or her. Whether or not the value of one descriptive variable actually changes is determined by a Monte-Carlo process. Let us consider a stylised example: consider an individual of a certain age. Given this age, he or she faces a certain risk of mortality denoted by d. Now for our individual, a random number between 0 and 1 is drawn from an uniform distribution. If the resulting number is below the mortality risk d (which can in turn be a function of other variables) then the individual is considered dead at $t+1$. If not, he remains alive, with the result that his age is increased with one. Likewise, our living individual faces a certain probability (technically speaking) of becoming married, having a child, finding or losing a job, and so forth. The number of variables which can be altered between subsequent points in time depends entirely on how much information on transition probabilities is available to the constructor of the model.

With dynamic microsimulation, the life history as well as earnings history of individuals belonging to different groups within the sample can be simulated. The modeller has relative freedom in what to add to the model. An important advantage is that it is possible to define individual stock-variables, adding up past values of the flow-type. For instance, the lifetime-income of an individual can be kept track of by adding up past (discounted) annual values of income over the whole lifetime of the individual. As a result, the effect of socio-economic policy measures can be expressed

[5] Indeed, this distinction between 'static' and 'dynamic' microsimulation models is somewhat confusing, since both types of models are dynamic in the sense that they are time-dependent. So, a formally better (but less appealing) description could have been 'statically time-dependent' and 'dynamically time-dependent'. The reader should keep in mind that both static - and dynamic

in terms of 'lifetime-income' instead of annual income of various groups and generations of individuals, since the latter can be expected to be biased in the sense that the redistributive effect of a certain measure is generally overestimated when expressed in terms of annual income. (Nelissen, 1995, and Harding, 1993). As will become clear when discussing static microsimulation models, this is one of the more important advantages of dynamic- over static microsimulation models. However, dynamic microsimulation models have disadvantages too: first of all, they are generally very large in terms of source code, they usually are very complex and take a long time to develop. As a result the costs of maintenance are high, and it takes quite some time to introduce new researchers to the technical details of the model. Moreover, their use and trustworthiness is restricted to the availability of trustworthy transition data. Lastly, dynamic microsimulation models make extensive use of computer recourses, though this is becoming less important due to the rapid development of computers. Moreover, as opposed to static models, dynamic microsimulation models do not allow the immediate jump from today to -say- 2020 without having to simulate all the intermediate years. An application of dynamic microsimulation on the pension system in the Netherlands is Dekkers et. al, 1995.

1.4.2. Static Microsimulation.

In her 1993-book on microsimulation, Harding (Harding, 1993, p.19, see also Harding, 1996, p.3) describes two key techniques involved in the static ageing of a dataset. The first one is to *reweigh* the sample, whereas the second key technique is referred to as *uprating*. Both techniques are used in STATION, the static model under consideration here. However, as it is the most fundamental technique of the two - and very typical for static microsimulation models, this section will concentrate on the first key technique: the reweighting of the dataset.

As said in paragraph 1.3, the reweighting-technique of static microsimulation is basically cold-deck imputation of missing variables. The vast majority of cross-sectional datasets contain a weighting variable which gives the individual more or less importance (i.e. a greater or smaller weight) in the sample in order to make the sample

---

microsimulation models include a notion of 'time', being reflected in changing (demographic) circumstances.

more representative for the whole population, for instance by neutralising the effect of selective nonresponse. The technique of static microsimulation boils down to adjusting these individual weights to let the dataset in the base year *t* meet descriptions of the future population, which are exogenous from the point of view of the model.

Suppose for instance that a 1992-dataset consists of a certain percentage of female individuals aged between 15 and 19 and suppose that we know from demographic projections that this proportion will decrease by 8.5% between 1992 and 2020. Then the '2020-proportion' of women can be formed by multiplying the weight variable of the women in this age group by (1-0.085)=0.915. Note that the weight factors of other categories must be adjusted upwards to neutralise the effect of this decreasing proportion of young women on the weighted size of the dataset as a whole.

The basic difference between static and dynamic microsimulation models is that the actual individual data remain unchanged in the case of static microsimulation modelling; only the weight factor is altered corresponding to the future situation. In dynamic models, by contrast, the weight variable remains unchanged but the actual individual information is changed according to individual transition risks and using a Monte-Carlo process. In reality, however, the difference is often less clear, since both techniques can be used in the same model.

The disadvantages of dynamic models are the advantages of static microsimulation models: the latter are technically simple (relative to dynamic models, that is), though less intuitive and less CPU-demanding[6] than dynamic models. This efficiency is increased further by the fact that, one can form the '2020-dataset' in one step, without having to simulate all the intermediate years first. Moreover, one can use the simulation results in other empirical research without having any technical knowledge on the model itself. In the case of STATION, the simulated weight-transformators can be downloaded from the homepage.

A drawback of static models is its lack of flexibility, compared to dynamic microsimulation models. Moreover -and this is probably the most serious drawback of static microsimulation models- the fact that simulation periods in time can be skipped reveals that the model does not have a 'memory'. To make this more clear, let us

[6] Of course, that is a relative notion: on a 100 mhz. Pentium with just 32 mb. RAM, running all the modules of STATION for all years between 1992 and 2050 takes at least three days and four nights.

return to the example of how to calculate the lifetime income (i.e. the income which an individual earns over his or hers whole life) in a dynamic microsimulation model, and consider whether or not this could also be done in a static microsimulation model. In order to construct the lifetime income of an individual, the annual income of this individual in each year of his or her life is added up. So, a new variable is created when an individual is born and increased each year, until the individual deceases . Is this possible in a static model? No, simply because the individual does not get born, gets older and deceases in a static model. The individual data remains unchanged: only the weights change. Moreover -again as opposed to dynamic microsimulation- the simulation results for a certain year *x* do not influence the simulation results of another year *y*, since these simulation results are both directly calculated from the base-data set. Consequently, the adding up of annual income -even if it would be meaningful- is not possible.

This discussion of the drawbacks of static microsimulation models ends the first part of this text. In this part, an introductory overview of microsimulation and the two types of microsimulation models was given. Both static and dynamic microsimulation models have their advantages and disadvantages, which are to a certain extent mutually exclusive. In the second part of the text, which now follows, the static microsimulation model STATION, developed by the author at the Centre for Social Policy (*Centrum voor Sociaal Beleid* or CSB) of Antwerp University (UFSIA), will be presented and discussed.

Chapter 2. Antwerp STATION.

When the need for a microsimulation model emerged, there was consensus among the CSB-researchers that this model should meet a number of demands, of which a short development period was not the least important one. Another thing was that there was doubt about the availability of enough transition data for a dynamic model. Moreover, a crucial demand was that the model should make it possible to be linked to other models of the CSB.

For these and other reasons, it was decided to build a static microsimulation model, named STATION (from STAtic microsimulaTION, indeed: the name of the model is also inspired by the fact that the author spends several hours in the train every day to commute between home and work). STATION describes the future development of the Belgian population starting in 1992 and allowing simulation of every year up to 2050. It aims at analysing the effects of demographic change (notably greying) on the Belgian social security system, and it is designed to be a complementary part of MISIM, the microsimulation model describing the Belgian social security- and tax-system, developed at the Centre for Social Policy. This model MISIM (MIcroSImulatieModel) is a static microsimulation model for direct taxes and benefits[7]. The relation between MISIM and STATION is best described by quoting Merz: '*Static microsimulation naturally is connected with the time period of the cross-section data* [which, in the case of MISIM is 1992, G.D.] *Temporal extrapolation to actualize the data or to forecast the sample into the future, called ageing of the sample, however, is available in more recent static* [microsimulation models, G.D.]' (Merz, 1994, p.6). It is this last sentence which describes the role of STATION.

The model STATION is written in SAS and consists of one general program and various submodules, which take the form of SAS-macros, generally with the projection-year as the only argument. The first submodule modifies the age- and gender structure of the 1992-dataset. The second submodule changes the distribution of the family-type to the future situation. This body of the model is completed by a number of SAS-macros, one of which deals with the intrapolation of the exogenous projection data. Other macros upgrade variables which reflecting economic growth, indexation and such. Moreover, additional and separate macros derive a number of dependent informative variables, such as poverty rates, income inequalities and so forth.

The model STATION transforms weight variables, which are then applied to the 1992-wave of the Socio-Economic Panel of the Centre for Social Policy. The SEP-panel data set started in 1985 and continued in 1988, 1992 and recently 1997. The 1992-wave of the SEP consists of 3821 households, of which 2285 are Flemish and 1177 are Walloon (Cantillon, et. al., 1993, p.7). Due to its size and as a result of weighting techniques, aiming to correct for selective attrition and non-response, the panel can be considered representative for the population (idem, p.7 and Proost, et.al.,

[7] Verbist, et. al., 1999, for an application of MISIM see, for instance, De Lathouwer L., in: Harding (ed.), 1996, Amsterdam, Elsevier, p. 69-91.

1996). So, every household in the dataset has an accompanying weighting variable. The dataset consists of a number of individual- and household characteristics. Monthly available income (of which the most important are net labour income and various sources of social security income) is gathered on the household level as well as on the individual level. Moreover, some household-specific income sources (such as housing grants) are also asked for. Lastly, and less relevant in this study, there is an extensive list of individual's consumption pattern, their attitude towards their income, and so forth.

The simulation results of the core modules of the model STATION are basically growth rates, which can be used to transform the household population weighting variable of the Socio-Economic Panel. These latter results will be presented in the third chapter of this study. The transformation of sample weights can be done on the individual - or the household level. For any future year between 1992 and 2050, the model generates a list containing the following variables for every individual.

1. **INDnum:** an unique number for every individual in the 1992-dataset.

2. **LIPRO:** family-classification of individual INDnum.

3. **weegL:** individual transformation according to the future age-distribution (module 1).

4. **weegG:** individual transformation according to the future LIPRO-distribution (module 2).

For every future year and for every individual in the 1992-dataset, the two transformation variables are generated by the model.

How can these simulation results be used? Suppose for instance that one wants to see how the income distribution of a certain subset of the Socio-Economic Panel changes between 1992 and, say, 2030. Or -which is possible as well- suppose that one wants to know how the estimation results of a certain behavioural relation change between 1992 and 2030. How can this be done? First of all, one derives the unique individual identification number INDnum out of other variables in the SEP-dataset[8]. Next, the above-mentioned list of transformation variables for the year 2030 must be combined with the SEP-dataset, using INDnum as the merging-variable. Thirdly, the household-weighting factor must be multiplied with one (or more) of the

[8] To see how this is done exactly, see the internet-homepage. One however needs the passwords for this.

transformation variables, depending on whether the variable is on the individual- or household level. The resulting dataset is the 1992-dataset, but then transformed to 2030 which can then be used to answer the questions stated above. If one wants to -as will be done in this study- one can uprate any monetary variable in the weighted dataset, using whatever assumptions one wants, provided that one does not want to use individual stock-variables, as explained earlier. In the presentation of the simulation results in paragraphs 2.1 and 2.2, uprating will be ignored, since the effect of reweighting will be shown by looking at some key demographic variables. In paragraph 2.3, the upgrading technique will be discussed in depth. The third chapter will entirely be devoted to the presentation and discussion of simulation results which combine reweighting and upgrading.

Before considering the model as well as the simulation results which stem from the weight-transformation process and upgrading, a final note must be made on the simulation years for simulation results will be presented further in this study: even though this will only be presented for the years 1995, 2000, 2005 and so forth, up to 2050, the model is capable to simulate all in-between years as well.

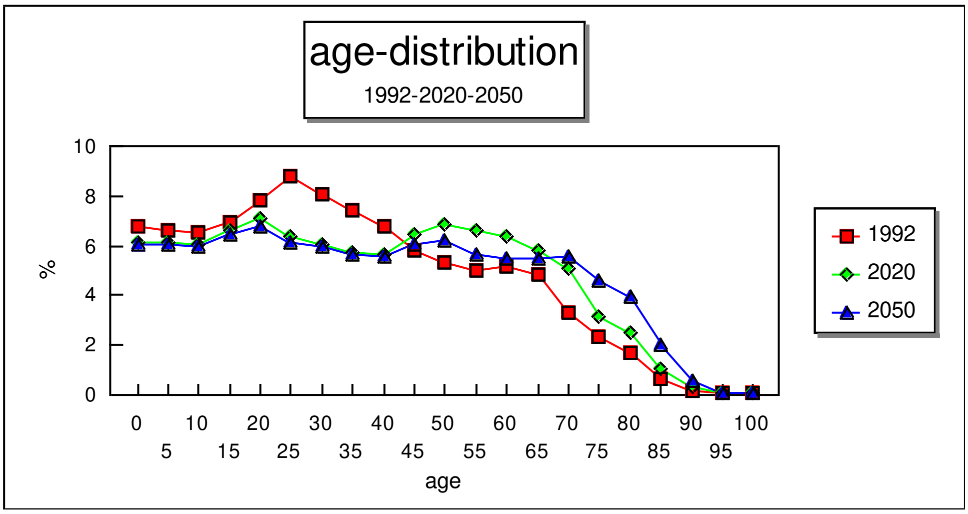
Next , the two modules which form STATION will be discussed in more depth. To see their effect on the data, figures describing the situation in the original dataset of 1992 will be compared to simulation results for the years 1995, 2000, 2005 and up to 2050. These simulation results will only be the result of the transformation of individual sample weights in the dataset and corresponding changes in income variables will therefore not be presented yet, since they lack realism.

2.1. The first key technique: reweighting to incorporate ageing.

The first submodule adjusts the age-gender structure of the 1992-dataset to the combined age-gender projections provided by the Belgian National Institute for Statistics and the Federal Planning Bureau (source: basic scenario of the *Bevolkingsvooruitzichten 1992-2050*). These projections show the future proportion of every age group - and gender combination in the dataset, relative to the 1992-proportions. The projections range from 1995 up to 2050 with five-year jumps (e.g. 1995, 2000, 2005, 2010...) and are indices with 1992 as the base year (or equal to 100). However, in order to make the model able to simulate to intermediate years as well, these projections were intrapolated. Thus, the first submodule ables us to modify the age-gender structure of the 1992 dataset to every future year between 1995 and 2050. This modification is done by simply multiplying the weight factors with the corresponding future indexes.

As the Belgian population ages, one of the general effects of this first module is that the weighting factor of older individuals increases ceteris paribus, whereas the weighting factor of younger individuals decreases, as could be expected. This can be seen by looking at figure 3.

Figure 3: Age-distribution.



Ageing is caused by two independent effects. First of all, if a cohort is much larger than the succeeding cohorts, this cohort will have a disproportional influence on the age-structure of the entire population. This is sometimes compared with a piglet swallowed by a snake (Becker, 1994, p.212, Lesthaeghe et. al, 1998, p. 61). After being swallowed, the piglet is pushed through the snake, while slowly being digested. Analogous to this, the babyboom-cohort moves right along the horizontal axis, while its size decreases as a result of mortality. The second -and probably more important- cause of the ageing of the population is the extension of the life-expectancy, resulting from technological, medical and economical development. The effect of the longer life-expectancy is reflected by the decrease of the size of the cohorts, or the speed of digestion of the piglet by the snake. This effect does not come forward from the above graph very clearly, but it can still be seen that the difference in the size of the 'bump' decreases harder between 1992 and 2020 than between 2020 and 2050. Thus, as time goes by, the negative effect of the mortality rate on the size of the bump decreases, showing the effect of decreasing mortality rates.

It could seem from figure 3 that there is virtually no difference between the proportions of the oldest age-groups at the three points in time. Concluding that the relative size of these oldest groups does not change would be erroneous, since this change is surpressed by the scale of the y-axis. Figure 4 shows the average growth rate of the proportions of the age-gender groups over the whole projection period. Here, the scaling effect is neutralised and the important effect of ageing on the relative sizes of the oldest age-groups is easily seen.

Figure 4: Average growth rate of the age-gender proportions.

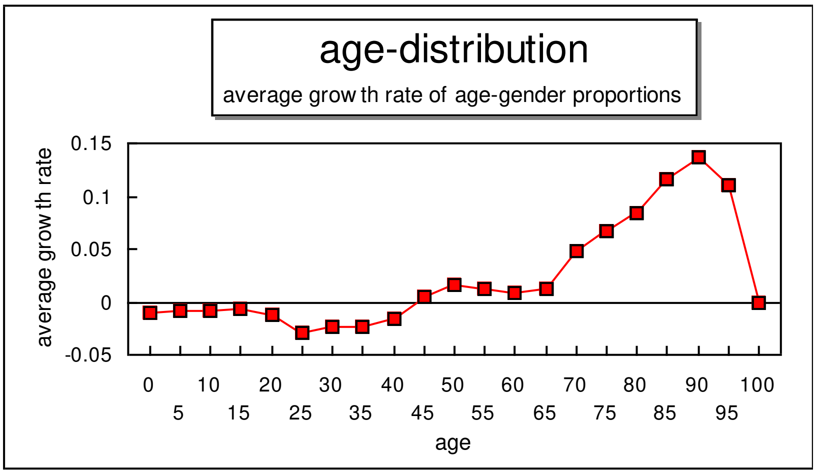


Figure 4 clearly depicts that, as a result of ageing, the size of the older age-groups will increase (especially the very-old), whereas the relative size of the younger cohorts decreases. The turning point lies around the age of 45.

2.2. The first key technique: reweighting to incorporate changing family structures.

In the second submodule, the family structure of the 1992- dataset is altered to meet family structure projections based on the 'realistic scenario' of Boulanger, P., A. Lambert, P. Deboosere and R. Lesthaeghe (*la Formation des Familles: étude Prospective*), of which the projection period is from 1996 to 2011[9]. Before proceeding

[9] As the entire projection period goes to 2050, for the simulation years after 2011, the household-transformation rates are based on the 2011-proportions. The implicit assumption is therefore that the family-structure given the age-structure does not change after 2011 anymore. Of course, this does not prevent the family-structure to change in the sample after 2011, but that would be due to a changing age-structure.

with the explanation of this module, it is important to realise that the distribution by the family-type LIPRO changes due to the individual ageing as described in paragraph 2.1. as well. For instance, as more older individuals live alone, the number of one-person households can be expected to increase. From Boulanger et.al., we know what percentage of the future population will be in what LIPRO-classification, given their age and gender. The model compares the percentage of the 'age-gender modified' sample with these projected proportions, and transforms the individual weighting factors to meet these proportions. Consequently, changes in the distribution of family-type which are independent of the age-gender distribution are imposed on the data. In other words, this second modules changes the distribution by LIPRO given the distribution of age and gender.

First of all, we need to know the different types of family which are distinguished. This classification is known as the LIPRO-classification and consists of the following entries:

1. Child of a married couple.
2. Child of an unmarried couple.
3. Child of an one-parent family.
4. Single.
5. Married individual without children.
6. Married individual with children.
7. Cohabiting individual without children.
8. Cohabiting individual with children.
9. Head of an one-parent family.
10. Living in the same house as 4, 5 or 6.
11. Others.

The last category (others) a.o. contains individuals living in nursery homes, psychiatric institutions and other (nonvoluntary) collective forms of cohabitation. However, these individuals do not occur in the 1992-dataset, so a direct comparison to the proportions in this dataset with the Boulanger-projections is not possible. To overcome this problem, the projected proportions have been recalculated, excluding this last category. Using the above proportions and the projected numbers of individuals in every age-gender-category, the projected numbers of individuals in each age-gender and family-type were derived. Next, the numbers of individuals in the last category ('others') were excluded from the dataset, and the proportions were calculated again.

In the following figures, the numbers of individuals in a certain family-category are represented as a fraction of the total number of individuals in that same age-category. In other words, these figures representing the probability that an individual is to be observed in a certain family-type, *given that he or she is in a certain age-category*. Taking the percentages to the total numbers of individuals in every age group largely neutralises the effect of ageing, since the first module transforms the weighting factors for age-groups as a whole. This means that in the following figures, the changes are mostly due to changing family-type distributions, and (to a large extend) disregard ageing.

Figure 5 & 6: LIPRO-classification, % age group.

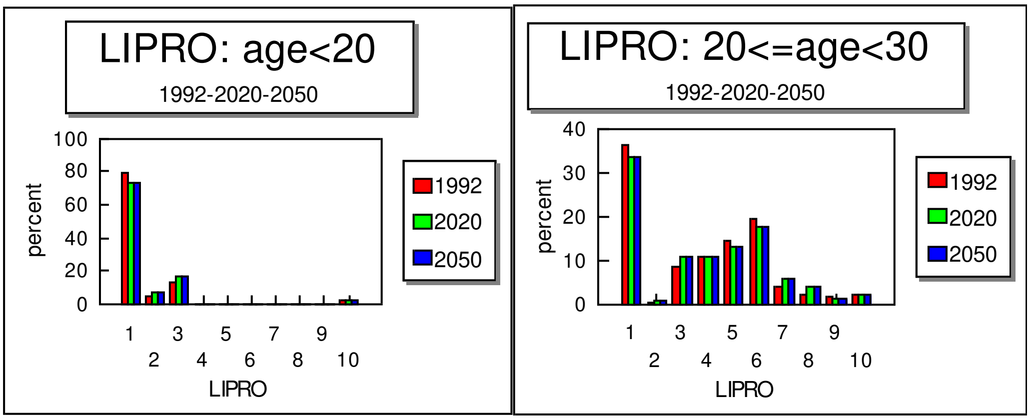


Figure 7 & 8: LIPRO-classification, % age group.

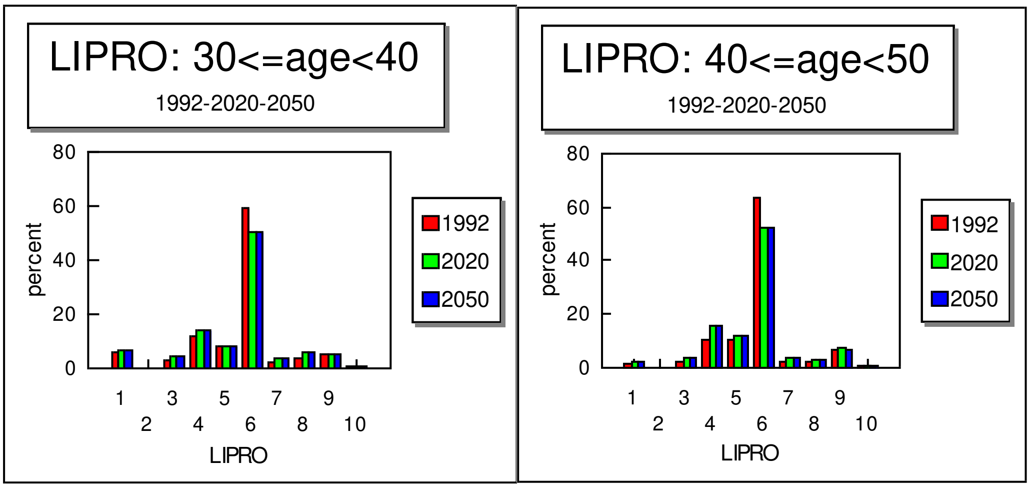
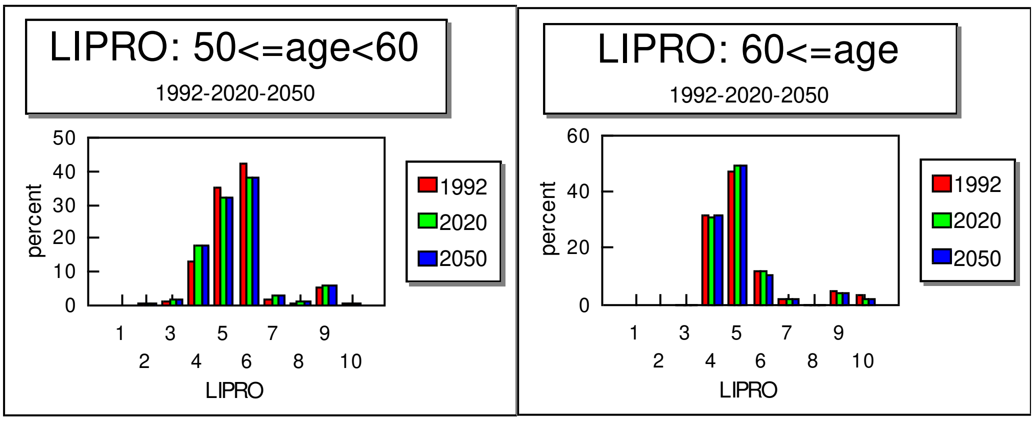


Figure 9 & 10: LIPRO-classification, % age group.



Remember that the horizontal axis consists of the household-types, of which the entries are given on page 20. In figure 5, it seems hardly surprising that the vast majority of the individuals younger than 20 are children of married couples (1), as this is the most traditional family-form. It is not very surprising either that the importance of this group is decreasing over time, a decrease which primary occurs in the first half of the simulation period. It is quite interesting to see that the relative number of children living in one-parent families (3) is higher than children from unmarried couples (2). The importance of both categories increases as the relative size of category 1 decreases. Again, this increase is the most important between 1992 and 2020.

For the group of individuals between 20 and 30, as shown in figure 6, the most striking is that -even though the relative sizes of individuals in other categories have increased relative to figure 5, the majority of the individuals in that age group still live with both their parents (1). Note that the relative number of individuals being a child of an unmarried couple (2) has decreased even relative to the relative number of individuals being a child in an one-parent family (3), who has decreased as well. A bit over 10% of the individuals between 20 and 30 are single, a percentage which is lower than the relative number of individuals which are married with -or without children (6 and 5). Especially the group of individuals belonging to the 6th category (married with children) is quite large. The relative sizes of the remaining categories are rather low, where it is noticeable that the group cohabiting individuals without children (7) is larger than the group cohabiting individuals with children (8): apparently, the event of getting children seems to be a motive for marriage.

When we look at the intertemporal shifts in figure 6, we see that the relative number of children in two-parent families and married individuals decrease, whereas the relative number of cohabiting individuals increase and so does the relative number of children in one-parent families.

The relative number of married individuals without children (5) decreases between figures 6 and 7 (individuals between 20 and 30 years of age and between 30 and 40 years of age), but remains fairly constant between figures 7 and 8 (individuals between 40 and 50 years of age). Meanwhile, the relative number of singles (4) increases somewhat at first, but then remains constant.

The pattern of the increasing importance of the group of married individuals which children, continues for individuals between 50 and 60 years of age. However, for the individuals over 60, a reverse shift can be seen: as children move out of the house and as the mortality rate increases, the relative numbers of married individuals without children and singles increases whereas the relative number of married individuals with children drops significantly between figures 9 and 10.

2.3. the second key technique: uprating.

2.3.1. Introduction and assumptions.

As said, Harding (Harding, 1993, p.19) sees reweighting as one key technique involved in the static ageing of a dataset. It is certainly the technique which disentangles this type of models from other models, which are often time-series models. The basis of STATION is formed by this reweighting. The second key technique is uprating, *where attempts are made to adjust monetary values to account for movements since the time of the survey or future anticipated movements* (idem). This second technique will allow us to distinguish various simulation-variants.

Up to now, we implicitly assumed that the nominal wages or social-security benefits do not change over time. In other words, the macroeconomic world was implicitly assumed to be 'frozen' in 1992; only the population was altered. This is a very unrealistic assumption, of course. Not only do productivity changes cause wage changes, but pensions in Belgium are linked to the rate of inflation which means that they increase as well (though generally at a lower rate than the wages). Moreover, the level upon which the wages set the (future) pension benefits, is subject to change as well. So, in order to make the model more realistic, this reweighted dataset must be supplemented with assumptions about the macroeconomic context. But what is the relevant macroeconomic context and how will it change in the future? Before turning to the actual discussion of how the monetary values were uprated, let us first consider briefly the macroeconomic assumptions underlying this uprating process. In order to keep things simple and workable, we only take the real development of wages and social security benefits as the exogenous time-variables. Thus, we require assumptions on the future macroeconomic development of real wages and of the indexation process. The course of the pension incomes is legally linked to the rate of inflation and therefore does not follow the (real) wage growth. Moreover, as the model is in real terms, the requirement that the pensions are linked to the price-index in this context means that they remain constant over time, whereas wages show a certain annual increase. As the projections of the Federal Planning Bureau of Belgium form the scientific point of departure of this study, let us remain as close as possible to the assumptions made in this study. This point of departure is an assumed real wage-increase of 2.25% per year (see Festjens, 1997, page 26) which is based on the average adjustments in the period 1969-1991 (idem, p. 82). So, as a **base-rate simulation or first scenario**, we should consider the distributional effects - and the relative income of the retirees given a constant real pension income and a real wage growth of 2.25% per year. This first scenario reflects what is expect to happen if no additional policy measures will be taken, i.e. if the current situation will persist. A **second scenario or - simulation variant** concerns introducing another assumption of the Federal Planning Bureau, namely a limited linkage of the pension benefit to the real course of wages. More specifically, we consider a annual increase of the pension benefit of 1%, given the 2.25% wage growth per year (see Festjens, 1997, page 33). This assumption is claimed not to be pessimistic, and for the real growth rate of wages, we most certainly agree, not to mention that it could even be considered somewhat optimistic. However, the distributional effect of the partial linkage of (pension) benefits to the course of wages remains un- or underexplored in the study of the Federal Planning Bureau. The **third scenario or - simulation variant**, a so-called 'wage-barrier' will be introduced. The purpose of this barrier is to limit the limit the pension benefit of all non-civil servants to a certain maximum, and -which is more relevant in the context of our model- to limit the effect of wage-increases on (future) pension benefit for those who are not civil servants[10] and who earn more than the wage-barrier. Again, following the assumptions of the Federal Planning Bureau, it will be assumed that this wage-barrier has a real growth rate of 1%, starting in 1997[11] (see Festjens, 1997, page 32). Lastly, the **fourth scenario or - simulation variant** will simultaneously introduce the partial linkage between pensions and wages, and the wage-barrier.

2.3.2. The link between wages and pensions.

Having described the various scenarios to be dealt with, let us describe these in more detail, since understanding of what is happening is crucial for the understanding of the simulation results. First of all, let us concentrate on the first scenario. The question therefore is: what is the effect of the assumptions of no linkage between the wage-rate and the pension benefit on the distribution of income, the income-inequality and the poverty among the retirees, given an assumed wage-growth rate of 2.25% and the demographic changes described in the earlier chapters? To answer this question, the first thing we have to know is how this process of linking the incomes of the retirees to those of the non-retirees is modelled. This is kept as simple as possible and uses the fact that, given an annual increase of the wage-rate of 2.25%, if an individual

[10] Those employed by the government, but on a temporary basis are not considered to be civil servants. For civil servants, the pensions are liked to the course of wages. This is the so-called "automatische perekwatie".
[11] Festjens uses this constant growth rate from 1999.

becomes 65 one year after 1992[12], his or her pension income can be expected to be 2.25% higher than the pension benefit of someone who retires in 1992. Now suppose an individual who is 67 in the year 1995, or 3 years after the base-year 1992. Then we know that he was 67-3=64 years of age in 1992. Using the above line of reasoning, we then know that, compared to the pension benefit of someone who is 67 in 1992, his or her pension benefit has increased with $(1+0.0225)^{65-64}$, accounting for the one year between 1992 and 1995 that the individual was not retired, and with $(1+L\times0.0225)^{67-65}$ for the two years between 1995 and 1992 that he or she was retired. In this, the variable *L* denotes the 'rate of linkage' which is between 0 and 1 (in the case of no link at all and a full link between wages and pension benefits, respectively). More generally speaking, the difference between the average pension benefit of somebody who is *a*>65 years old in the future year *y* and in 1992, can be written as

$$(1+0.0225)^{\alpha}\times(1+L\times0.0225)^{\max[0,(y-1992)-\alpha]}$$

Equation 1

where $\alpha = \max[0,(y-1992)-(a-65)]$.

The advantage of this upgrading technique is that it is simple and straightforward. The disadvantage should be mentioned as well, however, and is that the model does not allow 'backward-changes'. If we would set *L* equal to zero in the year 2000, for instance, the model would uprate the pension benefits as if there was no indexation from 1992 up to 2000. If we would then set L equal to .5 in 2005, the pension benefits would be uprated as if there was 50%-wage indexation from 1992 up to 2005. In other words, no past values of L are taken into account. The user of the model therefore has compete freedom in picking *L* for any simulation year, but the model will behave as if this value of *L* has been set at this particular value since the beginning of the simulation period i.e. since 1992. Of course, this is the direct result of the fact that the model is not capable of forming individual stock-variables, as mentioned earlier. The model disentangles various 'upgrading-regimes' for the various individual income-components in the dataset: labour income increases with 2.25% per year, whereas pension income remains constant (in real terms).

What will happen to poverty rates and income inequality in this first scenario or - simulation variant, i.e. if we let wages and (non-pension) social security benefits

[12] Indeed, the implicit assumption is that all those who are not yet retired, will retire at 65.

grow with 2.25% per year while leaving pension benefits unchanged (*L=0*)? It is clear that this will result in higher poverty rates, as compared to the situation where wages do not change as well. Likewise, income inequality can be expected to increase as well, as one category of individuals (namely pensioners) will see their level of their income deteriorate, relative to that of non-pensioners.

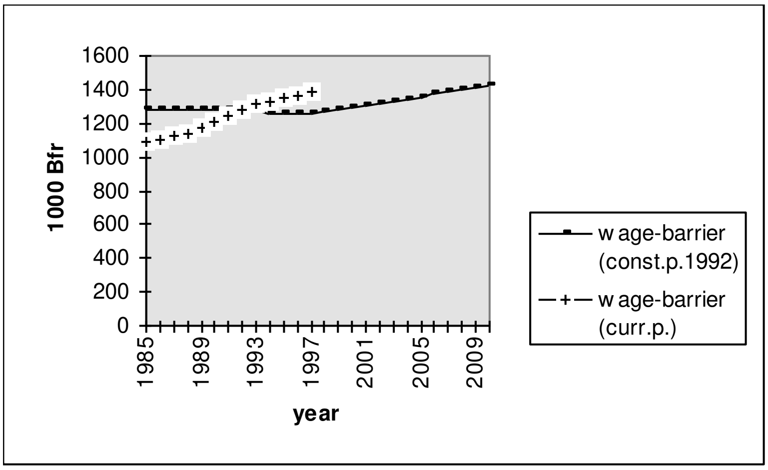
The next important question which will be dealt with, is the following: what is the effect of the assumption on partial indexation which the Federal Planning Bureau of Belgium uses as a basis for their projections of the Belgian state-wide pension system? Remember that these assumptions were an annual wage growth of 2.25% and an 50%-indexation (*L*=.5). This is the second scenario or - simulation variant. Compared to the first variant (economic growth and no linkage to the wage-rate whatsoever), we can expect the relative income position of the retirees to deteriorate less than in the case of no wage-indexation, but more than in the case of no growth of wages (the base-variant).

2.3.3. Introducing a wage-barrier in the determination of the level of the pension benefit.

The third and fourth simulation variants both introduce another additional assumption, which is that the future pension benefits of individuals who are not yet retired and who are not civil-servants, are subject to a ‘wage-barrier’. Since this wage-barrier is a new concept which has not been mentioned before, it requires some explanation: since 1981, employees contribute a certain fraction of their full income to the pension system. By contrast, the benefit which they can expect to receive is determined by the wage*, up to a certain maximum*. So, if an employee earns more than this wage-barrier, one contributes pension premium over the entire income, whereas the future pension benefit is limited to the income below the wage-barrier. Moreover, and that is relevant in this context, if an individual earns more than this wage-barrier, the future pension benefit does not follow the 2.25% annual increase of the wage-rate, but only the annual increase of this barrier. In other words, this wage-barrier makes the pension contributions progressive, implying some solidarity among pension-contributors. Moreover, as the growth rate of the wage-barrier is lower than that of the wages themselves, it causes the distribution of pension benefits to become smaller over time. The Belgian Federal Planning Bureau assumes that the wage-barrier increases only by 1% per year, whereas wages increase by 2.25% per year. The result of this is that more and more working employees will find themselves with a gross income up or above the wage-barrier. As a consequence, the increase of the average pension benefit will over time slow down, compared to the growth rate of the average wage. The problem which will be dealt with here is what the effect of this assumption is on the distribution of pension income, income inequality and so forth. The third simulation variant partially introduces the wage-barrier without assuming partial linkage between the course of wages and the pension benefit. The fourth and last scenario combines the assumptions of scenarios 2 and 3: it introduces a simultaneous partial linkage between pension benefits and wages and the implementation of the wage-barrier.

How does the model include this wage-barrier? This will be explained shortly. But first of all, what is the historical course of this wage-barrier? Figure 11 depicts the historical course of the wage barrier, both in current prices and constant prices (1992=100) where the latter is completed with the fictitious future development given a growth rate of 1% per year.

Figure 11: development of the historical and fictitious wage-barrier (source: van Eeckhoutte, 1997, p. 299).



The wage-barrier in current prices is almost 1.1 million Belgian francs per year (Bef. 1.089.988, to be exact) in 1985. It then increases to almost 1.4 million francs in 1997. However, if expressed in constant prices of 1992, this increase is reversed in a decrease, though of a very small magnitude. It seems that the actual implementation of a 1% real annual growth of this wage-barrier would imply a change in the trend of the last decennium.

If STATION would have been a dynamic microsimulation model, the inclusion of an increasing wage-barrier would have been very simple. For every employee, we would have compared his or her gross labour income -being the base of the future pension income- with the wage-barrier. Then -if necessary- the future pension income would be set as a function of the labour income, but only up to the wage-barrier. However, in a static microsimulation model such as STATION, individuals do not 'shift trough time'; only the weight factors change. Consequently, the above technique can not be applied and the inclusion of a wage-barrier is less straightforward, although still possible. The basic line of thought lies close to the way in which the indexation process is modelled and uses the fact that, given an annual increase of the wage-rate of 2.25%, if an individual becomes 65 retires one year after 1992, his or her pension income can be expected to be 2.25% higher than the pension benefit of someone who retires in 1992, that is*, if one did not earn more than the wage-barrier,* of course. If so, this effect of wage-change is limited to the effect of the change of the wage-barrier. As said, this change is based on the historical figures from 1992 to 1997 (which means it is negative but very close to zero) and assumed to be 1% from 2000 onward.

The implementation of such a wage-barrier involves several serial steps: first of all, given the assumptions on the level of the wage-barrier and the annual growth rates of both this wage-barrier and the general wages, calculate the proportion of employees and self-employed aged between 60 and 64 who will have a gross income up or over this wage-barrier in the future year y. Call this percentage p_y. This number p_y can be interpreted as the probability that one will draw an older employee whose income is limited in its effect on the pension benefit from the sample of older employees. The reason why only older employees are selected, is that the assumption is that the 'profile' of older employees can be used as a proxy for the profile of younger retirees[13]. The second step then involves the multiplication of the number of

[13] Indeed, by doing so, a certain overestimation of the empirical effect of the wage-barrier on pension income is inevitable and must be accepted. The reason for this is the following: if the proportion of workers earning more than the wage-barrier increases in the future, there is reason to believe that it has been doing so in the past as well. So, in a certain year, the proportion of young retirees who have earned more than the barrier in their last years of their career, can be expected to be lower than the proportion

retirees in the year y who have reached 65 between 1992 and y and who have been either employees or self-employed (the group of potential limited retirees) with this p_y, resulting in the number of 'actually' limited retirees. Denote this *LR*. So we are now in the situation that *LR* individuals must be selected from the group of retirees which meet the above-stated requirements. How can this be done? By simply using the fact that only the richer pensioners (those with the highest pension benefit) will have 'encountered' the wage-barrier. So, serially rank the sample according descending pension income and select the first *LR* individuals from this ranked set of data. For these selected individuals, the uprating-equation becomes

$$[1+(1-WB)\times 0.0225]^{\alpha}\times(1+L\times 0.0225)^{\max[0,(y-1992)-\alpha]}$$

Equation 2

where $\alpha = \max[0,(y\text{-}1992)\text{-}(a\text{-}65)]$

for somebody who is a>65 years old in the future year y. The variable 1-*WB* denotes the growth rate of the wage-barrier as a fraction of the growth rate of the wage. It is written as above so that *WB*=0 implies that the wage-barrier is fully linked to the real course of wages, which in the context of this uprating-model is equivalent to saying that the wage-barrier does not exist at all[14]. Following the assumption of the Federal Planning Bureau, the annual growth rate of the wage-barrier is 1%, whereas the growth rate of wages is 2.25%. *WB* is therefore equal to 1-(1%/2.25%)=1-0.444=0.556.

Now how does the introduction of this wage-barrier change the above results? To recapitulate, the introduction of a wage-barrier (of which the indexation is limited) implies that not all individuals see their pension income being fully indexed. This is only the case for those whose income is below the barrier or ceiling. Consequently, one can expect the poverty-increasing effect of wage-indexation of pension benefits to be reduced even further, compared to the case of limited or no- indexation, where this effect will only emerge in the longer run. And what can we expect to be the effect of

---

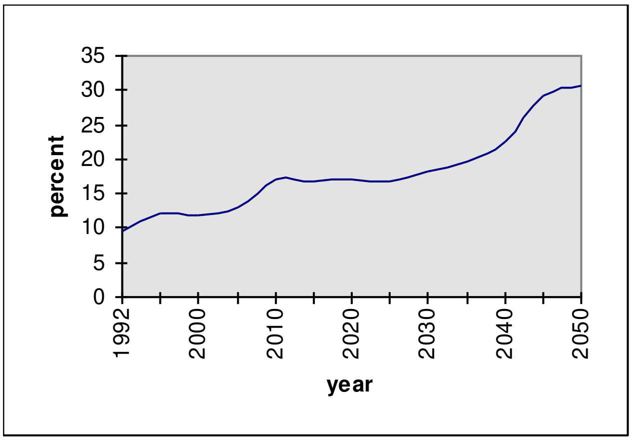
of current older workers who earn more than the barrier. However, trying to form a correction factor for this difference would be a very complex task (since current retirees have not all retired in the same year, but in one of a whole range of years) and would require quite some ad-hoc assumptions, which would make the model far more complex, and less trustworthy. Given the expected small magnitude of the overestimation (after all, the age differences between older workers and younger retirees are not that large since there are retirees younger than 65 as well) seems acceptable.

[14] In the future, that is. It implicitly does exist up to 1992 since its effects are reflected by the actual distribution of pension income in the data set.

the wage-barrier on the income distribution? Will income inequality decrease or increase? As a result of the implementation of the wage-barrier, the growth of the highest employees-pension benefits (and that of former self-employed) becomes dampened. As a consequence, one could conclude that these highest pension benefits would converge to the mean, thereby reducing the overall inequality of pension benefits. However, this conclusion would be wrong, since it implicitly assumes that the highest pension-benefits of ex-employees and self-employed are the highest pensions of the whole sample. But this might not be the case, since we could very well assume that the pension benefits of numerous former civil servants (which after all are based on the final wage instead of the average wage) will be higher than these limited pension benefits of ex-employees. As a result, the pension benefits which are limited by the introduction of the wage-barrier are not necessarily in the highest pension-income deciles and the effect of this wage-barrier on the distribution of income is therefore ambiguous. If the highest pensions of former employees are found high in the sample-wide distribution of pension benefits, then the introduction of the wage-barrier will result in a decreasing income inequality. But if, on the other hand, employees and self-employed are not so much found in the top of the income distribution (apart from a very small group of very high-income earners, maybe) then the implementation of the wage-barrier will cause income inequality to increase. In either way, this would only hold in the short run and noticing that the strength of this effect remains unclear. We said that this would be the case 'at least in the short run'. Doesn't this hold for the long run as well? No. In fact, whatever the short run effect of the implementation of the wage-barrier on the sample-wide inequality of pension benefits is, it can be argued that the inequality of pension benefits will increase in the very long run. To see why this is the case, let us recapitulate that the direction of the effect of the implementation of the wage-barrier on pension-income inequality depends on the location of the (limited) pensions of former employees and self-employed in the sample-wide distribution of income. If these pension benefits are high in the income-distribution, its limitation will cause income inequality to decrease. On the other hand, if these pension benefits are not in the top-percentiles of the pension-income distribution, the sample-wide income inequality will either remain stable or increase. Now let us combine the above line of reasoning with the notion of 'automatic perequation', as a result of which the pension benefits of the former civil servants are fully linked to the development of wages. So, irrespective of the simulation variant, the pension benefits of former civil servants increase by 2% per year. This in turn means that the pension benefits of former employees and self-employed deteriorate relative to the pension benefits of former civil servants, and this relative decrease becomes stronger the further in the future we go. Consequently, over time, the pension-incomes of non-civil servants will ‘move down’ the sample-wide distribution of pension incomes and the inequality-increasing effect of the implementation of the wage-barrier will therefore become stronger as time goes by.

Having said this, let us continue with looking at the actual data. First of all, how does the proportion of older employees (and self-employed) change as a result of the difference between the wage-change and the change of the wage-barrier? To see this, consider figure 12.

Figure 12: proportion of working employees and self-employed individuals between 60 and 64 who earn more than the wage-barrier.



According to the SEP-dataset, in the base-year 1992, 9.5% of the employees (and self-employed) earn more than the wage-barrier. This percentage remains somewhere around 12% until 2005, when it increases rapidly to a bit less than 17% in 2010. It then remains on the same level until 2025 when it starts a continuous increase, this time to 30.8% in 2050. It can easily be seen that it is not just the difference in growth rates between wages and the wage-barrier which causes the growth of the proportion of individuals earning more than the wage-barrier[15]; instead, the changing age-distribution clearly has a strong impact on this proportion[16]. As a last remark, note that this measure does not hold for civil servants.

Let us summarise the various scenarios by means of the following table.

Table 1: Simulation variants or - scenarios.

scenario	wage-growth	part. linkage of pension ben.	part. linkage of wage-barrier	Notation
first scenario or base-rate scenario	X	-	-	(n,n)
second scenario	X	X	-	(y,n)
third scenario	X	-	X	(n,y)
fourth scenario	X	X	X	(y,y)

[15] If so, the line would be convex and monotonically increasing.

[16] Showing the exact cause of these changes would ivolve knowing the probability of earning more than the wage-barrier for individuals of different one-year age groups and combining this with the changing age-patterns resulting from the reweighting process. This would lead us too far.

Table 1 clearly shows that all scenarios assume an annual real wage growth of 2.25% per year. The first scenario is the point of departure and assumes no linkage between wages and ongoing pensions (which means that pension benefits are only linked to the rate of inflation) and the non-existence of any wage-barrier (which is equivalent to saying that the wage-barrier completely follows the real course of wages, at least in the context of upgrading). The second and third scenarios introduce the partial linkage of pensions to the course of wages and the implementation of the wage-barrier, respectively. Finally, the fourth scenario combines the last-mentioned assumptions of the second and third scenarios. The notation will be explained later and is only mentioned for completeness' sake. Before ending this discussion of the simulation variants, and the potential policy measures which disentangle them, the reader should realise that these measures only describe the 'future', which in this model starts in 1993. For, the past situation and the policy measures then taken, are reflected in the 1992-dataset, which is transformed to describe the future situation. So, these past measures and their effect on the income distribution in 1992 are taken to the future as well. They can never be neutralised or changed.

2.3.4. Contributions.

Next, we turn to the income-side of the model. Hitherto, changes in total pension benefits paid out, did not have any consequences for the net-labour incomes. This of course violates the Pay-As-You-Go-character of the Belgian pension system. Of course, a changing total pension benefit, be it the result of a changing dependency ratio or resulting from a changing average pension benefit of the retirees, must be covered by a decrease of the total income of the non-retirees. As we deal with total household incomes, and to keep things simple, let us start from the assumption that every increase of the total income of the retirees in a certain year must be covered by the non-retirees in the same year. Denote variables expressed in capitals to be aggregates of individual-level variables. Then the following PAYG-equation holds for all years *t*:

$$\chi^t i_{nr}^t N_{nr}^t = \beta^t i_r^t N_r^t$$

Equation 3

where i^t_{nr} and i^t_r are the per-capita household incomes of individuals who are not retired, respectively retired in year *t*. Likewise N^t_{nr} and N^t_r are the numbers of retired and non-retired individuals in period *t*. Furthermore, denote χ^t and β^t the contribution- respectively benefit rate (i.e. the contribution, respectively benefit as a fraction of the wage) in time *t*. Suppose $\alpha^t = \chi^t / \beta^t$ so α^t is the contribution rate relative to the benefit rate. By definition, equation 3 must hold for period 1992 and all other periods *1992+z* ≤ 2050. So, the PAYG-equations for 1992 and *1992+z*

are $\alpha^{1992} i_{nr}^{1992} N_{nr}^{1992} = i_r^{1992} N_r^{1992}$ and $\alpha^{1992+z} i_{nr}^{1992+z} N_{nr}^{1992+z} = i_r^{1992+z} N_r^{1992+z}$. Of course, in practice, the above PAYG-equation does not hold for 1992. Total household income of the retirees is not generated by contributions of the non-retirees; only *pension income* is. However, formulating the problem the above way, gives the possibility to take the ratio of both the left and right-hand side of the PAYG-equation in both periods of time.

$$\frac{\alpha^{1992}}{\alpha^{1992+z}} \times \frac{i_{nr}^{1992} N_{nr}^{1992}}{i_{nr}^{1992+z} N_{nr}^{1992+z}} = \frac{i_r^{1992} N_r^{1992}}{i_r^{1992+z} N_r^{1992+z}}$$

Equation 4

Now the assumption that all 1992-income of the retirees is covered by contributions of the non-retirees no longer holds, since we have expressed the above equation in *relative changes*. Put differently, the implicit assumption underlying equation 4 is that the *relative change* of the income of the retirees, relative to that of the non-retirees, must be covered by contributions by the non-retirees. Without loss of generality, assume $\alpha^{1992}=1$ (so that α^{1992+z} denotes the change of the contribution relative to the unknown 1992- value) and rewrite equation 4 to:

$$\frac{i_{nr}^{1992} N_{nr}^{1992}}{i_{nr}^{1992+z} N_{nr}^{1992+z}} \times \frac{i_r^{1992+z} N_r^{1992+z}}{i_r^{1992} N_r^{1992}} = \alpha^{1992+z}$$

Equation 5

Equation 5 is expressed in such a way that the change of the contribution, α^{1992+z}, can be expressed as a combination of the inverse change of the total income of the non-retirees and the change of the total income of the retirees. By assuming the change of the per capita income is the same for both retirees and non-retirees (i.e. $i_{nr}^{1992+z} / i_{nr}^{1992} = i_r^{1992+z} / i_r^{1992}$) it can easily be seen that α^{1992+z} becomes a direct function of demographic ageing. Of course, we can also assume no demographic change (i.e. both the number of retirees and non-retirees increase by the same rate between 1992 and 1992+z) in which case α^{1992+z} becomes a function of the proportional change of the per capita incomes of the retirees and the non-retirees.

By multiplying the (1992+z)-household incomes of individuals by $1/\alpha$, we therefore take into account the fraction by which the contributions at t+z must increase. However, by doing so, it is assumed that all non-retired individuals contribute the same fraction of their income. Rather, it would be preferable to keep the progressivity of the coverage of this burden as close as possible to the progressivity of the tax system as a whole, since any real increase of the amount to be paid would be done through the tax-scheme. However, attempts to achieve this unfortunately failed, so it was decided to stick to the above model.

This chapter started with discussing extensively how the 1992- wave of the SEP-dataset is modified to meet future demographic changes, both on the age-distribution as well as the distribution of household-types given the age-distribution. The second step in static microsimulation is uprating. This was discussed in paragraph 2.3. This uprating takes the form of introducing a partial linkage of the pension benefit to the course of wages, as well as a possible implementation of a wage-barrier. As a consequence of both uprating and the transformation of the weights, a PAYG-equation was introduced in order to maintain the equality between contributions and benefits. In the next section, some practical problems which one gets to deal with when building a static microsimulation model, will be discussed.

2.4. Static microsimulation in practice: pitfalls and solutions.

In this section, some practical problems encountered when developing STATION will be discussed, together with the (sometimes more pragmatic than elegant) solutions to these problems. This section is meant for those readers interested in the practical 'how-and-why' of static microsimulation. However, having read it is not a necessary condition for the understanding of the other sections. Those who are just interested in the outcomes of the model and who thinks the general description given in the last paragraphs is enough, are invited to skip this paragraph and move to the third chapter, where the simulation results will be presented and described.

The first pitfall which will be dealt with in this section, concerns the so-called 'structural deviation of the weighting factors'. Next, the problems which come with observing empty cells in the 1992-dataset will be considered. Thirdly, problems related to sample bias in the original dataset will be considered. Lastly -and this is rather important- the fundamental problems involving the use of individual transformations when considering household-data will be discussed. As a result of these problems, we will adopt the household-average transformation for all individuals within a certain household.

2.4.1. The first pitfall: Structural deviation of the weighting factors.

First of all, let us consider the structural deviation of the weighting factors. There is no reason why the average age-gender index[17] to a certain future year, taken over all age groups for a certain future year, should be equal to 1. This is equivalent to saying that the average transformation of the individual weight variables need not be zero. On the contrary: one can expect the average index to be larger than 1 since this reflects the general population growth. However, the transformations derived in the second submodule are based on the difference between, on the one hand, the LIPRO-distribution of the age-gender modified dataset and, on the other hand, the projected LIPRO-distribution. As the percentages in both distributions necessarily add up to 100, the summed difference between the percentages (taken for each LIPRO-category) must be equal to zero and so should therefore the average difference. However, even though the summed difference between the percentages must add up to zero, there is no reason why the *growth rates* (or the transformation rates) should. To see this, consider the following example of two years and two categories. Suppose that in the base-year, the sample consists of two categories. Suppose that 40% of the sample belongs to the first category, which means that 60% belongs to the second category. Suppose furthermore that we know that in the future year, these percentages will be 50% and 50%, respectively. As the percentages add up to 100%, it is easy to see that the differences add up to zero (40%-50%+60%-50% = 0). However, the growth rates are .25 and -.166 and they clearly not neutralise each other. Fortunately, in practice,

[17] The number by which the household-weighting variable of 1992 must be multiplied to get the future weighting variable.

this does not turn out to be a major problem, since the percentages are not that deviant as they are in the above example. In practice, the 'structural deviation' (the difference between the actual average transformation and zero) is very small, fortunately: it does not exceed 2% for any future year.

What are the consequences of this structural deviation on the simulation results? If the transformations to family-classifications do not meet the requirement that they are on average equal to zero, this transformation affects the age-gender structure and even the size of the population in an unpredictable way. In other words, the transformation to family-classification are no longer neutral since they in some way affect the distribution of other variables, such as age or income, even if this nonneutrality is as small as 2%. To solve this problem, a two-step procedure is adopted. In the first step, the transformation rates for all combinations of gender-age group and family-type are calculated for a certain future year. These transformation rates are based on the (modified) projections of Boulanger et.al., and are nonneutral for the reasons just explained. In the second step, these individual transformations are divided by their overall average, thus imposing their neutrality without affecting the relative differences between these transformations.

2.4.2. The second pitfall: Empty cells.

The second pitfall which is to be discussed now, deals with empty cells in the original dataset. A cell is defined as a certain combination of characteristics, in this case gender and family-type. A modification made to the projected data is that the weighting to gender and family-type is done in ten-year age groups instead of five-year age groups, for both males and females separately. Moreover, all individuals of 70 and older are taken together for every combination of gender and family-type. Of course, this makes the simulation results less accurate since not all information is used. Nevertheless, the reason why this has been done, touches directly at the heart of static microsimulation. In the following table, the relative numbers of individuals being in one of four age groups and the LIPRO-classification is presented for 1992. Now remember that the individual weighting factors are transformed in such a way that the transformed sample distribution meets the projections of Boulanger et.al. For every combination of age-group and LIPRO-category, this is done by taking the growth rate between the 1992-percentage and the projected percentage. And this is exactly where things go wrong: for instance, from the Boulanger-tables, we know that about .01% of the population between 70 and 74 will be cohabiting and will have children living in the same house (LIPRO-category number 8).

Table 2: relative number of individuals in age-and family categories, 1992.

LIPRO	60≤age<65	65≤age<70	70≤age<75	75≤age
1	.00	.00	.00	.00
2	.00	.00	.00	.00
3	.42	.00	.00	.00
4	16.46	23.68	34.19	53.18
5	54.17	56.32	48.5	29.43
6	21.30	14.03	7.71	2.42
7	1.98	1.54	2.7	.61
8	.35	.00	3.88	.00
9	4.67	3.3	.00	6.00
10	.65	1.12	3.00	8.36

Unfortunately, in our 1992-sample, this percentage is zero. Consequently, it becomes impossible to calculate the growth rate and this in turn means that there will be no way in which the transformation of the weighting factors can ever make the sample meet the future distribution. In this and other cases, the whole technique of static microsimulation breaks down, since one empty '1992-cell' implies that the proportional sizes of other 1992-cells are relatively larger than they would have been, would the cell not have been empty. In other words, all the other growth- or transformation rates will be different from what they would have been if there were no empty cells. Consequently, the existence of only a couple of empty cells (combined with not-empty future-values of that cell) can jeopardize the trustworthiness of the entire model. There is no way in which this problem can be solved, if it emerges, but its implications can be minimised or even completely avoided by defining the different cells as such that the number of empty cells with non-zero projections is minimised, which has been done[18]. With regard to LIPRO-classification 8 (cohabiting with children) and 9 (head of an one-parent family), it is clear that clustering the columns to two age-categories (younger than 70 and older than 70) solves the problem

[18] Indeed, if the 1992-cell is empty, and the future cell is empty as well, there is no problem since the growth rate can then be set equal to 1.

of empty cells[19]. The interpretation of this grouping together basically is that it is assumed that a change in one age-group over seventy with respect to another age group over seventy, does not significantly affect the distribution of family-type. Stated differently, the probability that an individual will be observed in a certain family-type, is assumed to be more or less the same for all age groups from 70 years of age on.

2.4.3. The third pitfall: Sample bias.

The third pitfall which will be considered is related to the existence of sample-bias in the 1992-dataset. As said before, the projections of the age-gender structure of the population take the form of indices with 1992 as their base year. This basically implies that the 1992-proportions are reweighted according to these indices. As a result, if there is some sample-bias in the dataset, it will not be 'solved' by the weighting process: all projections for all future years will suffer from the same sample bias. As opposed to this, the reweighting of the sample with respect to the gender-family type distribution is done by comparing the 1992-proportions of LIPRO with the projected proportions. Consequently, if there is sample bias in the base year which affects the 1992-proportions, the weighting process will take that into account, and the sample-bias will not emerge in the projected data[20]. The drawback of this is that the comparison of the sample between the base-year and either one of the future years is 'polluted' by the sample bias. An idea on the possible existence of sample bias was formed by looking at the 1996-transformation values. The assumption was that, given age and gender, the probability of being observed in a certain family-type remained unchanged (or changed only very little) between 1992 and 1996. Under this assumption and correcting for ageing, the difference between the age distributions in 1992 and 1996 would reflect the sample bias. Indeed, the results suggested the possible existence of sample-bias, though it is empirically not very important. The number of children from one-parent families and the number of singles are underestimated (at most -.5%), whereas the number of married individuals, both with and without children, is overestimated (at most +.4%). A possible explanation for the

[19] It can be seen from table 2 as well, that the clustering of some LIPRO-categories while keeping the age-groups constant would not have solved the problem.

[20] Of course, sample bias on variables other than those in the weighting process are not taken into account: they do not cease to exist.

underestimation of the number of singles in the SEP data of 1992, relative to the projections of Boulanger et.al., is the following: whereas the SEP data are ‘actual’ data -gathered by questionnaires, the Boulanger-projections are based on official figures of the Belgian population (Volkstellinggegevens), which more reflect the legal (tax-) situation. It is very well possible that surviving partners actually cohabit, whereas they claim to be living alone, this to prevent losing pension rights.

Even though this sample bias does seem to be rather small at first sight, it turned out to have some effect on other variables (pension income, for instance). The result of this bias is that the comparison of the average pension income between 1992 and 1996 is polluted. This problem was solved by multiplying the 1992-dataset with LIPRO-proportions from the *volkstellinggegevens 1991*[21], thereby neutralising the sample-bias.

2.4.4. The fourth pitfall: Households versus individuals[22].

Now we turn to what probably is the most fundamental (as well as complex) problem which we will deal with in this section. Up to now, the simulation results were based upon individually-transformed weighting factors. These transformations are based on the indices reflecting the proportional changes, specified to five-year age groups and gender. To explain the methodology in more detail, let us turn to a strongly stylised example.

period t	household 1:	1 individual of 20 years old
		1 individual of 60 years old
	household 2:	1 individual of 20 years old
		1 individual of 30 years old

N=4

[21] The period of investigation of the SEP-data of 1992 actually started in december 1991 and ended in april 1992. The figures from the *volkstelling* are of the beginning of 1991, so the intertemporal difference is less than one year.

[22] The author would like to thank Karel Van den Bosch for his help with this paragraph.

Suppose that the population is as follows in the future period t+n,:

period t+n	3 individuals of 20 years old
	1 individual of 30 years old
	2 individuals of 60 years old

N=6

where N is the population size. Note that we have no information on which household what individuals are in.

The first step is the derivation of the population indices to the three age groups (20,30 and 60). This is done by, for every age group, relating the absolute numbers of individuals in both time periods to each other.

age	index	growth rate of index (t=100)
20	100*(3/2)	.5
30	100*(1/1)	0
60	100*(2/1)	1.

Note that the average growth rate of the index is equal to the population growth between t and t+n. (.(5+1)/3=.5 = (6-4)/4) So, the individual data is transformed on the basis of the growth rate of the population indices for period t+n. However, this causes a problem, since it means that individuals in the same household are transformed differently. As an example, let us consider household 1: the weight of the youngest individual is increased with 50% whereas the weight of the older individual doubles (the increase is 100%). As a result, the notion of 'household' loses its meaning, since the one individual becomes much more important than the other. One could say that a half individual of age 30 and an individual of age 60 enter the household, which means that the original household no longer exists, of course.

So, at the one hand, most income figures (average income, income inequality and poverty) are based on the household-income. However, at the other hand, the original households themselves lose their meaning, due to a inconsistent transformation within the household. Of course, this is a very serious flaw which undermines the trustworthiness of the simulation results. How can we solve this problem? We need a consistent transformator i.e. a transformator which is the same for all individuals in the household (demand 1). Moreover, this consistent transformator should not disturb the proportional sizes of various subcategories in the sample (demand 2); at least, the disturbance should be small. Thirdly, this consistent transformator must meet the same demand as the individual transformator (the growth rate of the index) does, namely that its average value is equal to the population growth rate (demand 3) . In our above example, we have two households. Denote the unknown household-transformators *A* for the first household, and *B* for the second household. Moreover, we have 3 age-categories, so that our model is overidentified. However, it can be shown that there is one unique solution in this very simple case. The equations describing *A* and *B* are the following:

1. $1*A + 1*B = 3$ (age category 20)
2. $1*B = 1$ (age category 30)
3. $1*A = 2$ (age category 60)

So, we get the solution *A*=2 and *B*=1. To verify these results, let us restore the derive the future numbers of individuals in the three age-categories, based on these consistent transformators. Let us start with the group of individuals of 20 years old. In the original dataset, there is one 20-year old in the first household and one in the second household. So, the future (transformed) number of 20-year olds is equal to 1*(2)+1*1=3. For the other age-categories, these numbers become 1*1=1 for the 30-year olds, and 1*2=2 for the 60-year olds. Indeed, these numbers match the future population as presented earlier. To check demand 3, we need to see whether the average value of the growth rates derived from A and B, is equal to the population growth rate, (which is 6/4-1=0.5). It is easily seen that this is the case: the average growth rate is (B-1+A-1)/2 =1/2=.5.

Of course, the above example is oversimplified in the sense that there exists an unique solution. In fact, in a cross-sectional dataset of reasonable size, one can assume that there will be more households than categories of relevant variables. For example, let us consider the following -again highly stylised- dataset, consisting of 5 households (in 1992) and 3 age-categories.

		unknown transformator
period t		
household 1:	1 individual of 20 years old	A
	1 individual of 30 years old	
	1 individual of 30 years old	
household 2:	1 individual of 30 years old	B
	1 individual of 60 years old	
	1 individual of 60 years old	
household 3:	1 individual of 20 years old	C
	1 individual of 60 years old	
	1 individual of 60 years old	
household 4:	1 individual of 20 years old	D
	1 individual of 30 years old	
	1 individual of 60 years old	
household 5:	1 individual of 20 years old	E
	1 individual of 20 years old	
	1 individual of 30 years old	

Assume again that we know from external sources that the population is at t+n will be exactly the same as in time t, but with one household added to it:

household 6:	1 individual of 60 years old
	1 individual of 60 years old
	1 individual of 60 years old

So, analogous to the first example, we now are looking for 5 unknown variables A to E, having 3 age-categories.

The equations describing these unknown variables are the following

1A+1C+1D+2E = 5	(age category 20)
2A+1B+1D+1E = 5	(age category 30)
2B+2C+1D = 8	(age category 60)

As there are 5 unknown variables and 3 equations, we can conclude that the model is underidentified[23].

To recapitulate, the number of unknown variables we are looking for is equal to the number of households, whereas the number of equations is equal to the number of categories the data is grouped into. If we have 2 households and two age-categories, the model is identified, meaning that the consistent household-transformations can be derived. However, if we have 5 households and 3 age categories, the model becomes to a large extent unidentified.

When introducing this second stylised example, it was claimed that this situation was more in line with reality than the first example. Now let's see if this is true. We know that the individuals in our actual 1992 sample can be subdivided to gender (2 categories), household-type following the LIPRO-classification (9 categories) and five-year age group (20 categories). Then the total number of categories is 2*9*20=360. This is the total number of equations. Again, for every household consisting of the same numbers of household-members in the same combination of these three dimensions (gender, LIPRO and age-group), an uniform (consistent) transformator must be derived. So, if we assume that every household consists of a number of individuals, occupying an unique combination of the three relevant variables, then the number of unknown variables is equal to the number of households, which is 3821. It is of course possible that more households form the

[23] Following the Rank Condition (Gujarati, 1988, p. 588) the matrix of coefficients is:

$$\begin{bmatrix} 5 & -1 & & 1 & -1 & -2 \\ 5 & -2 & -1 & & -1 & -1 \\ 8 & & -2 & -2 & -1 & \end{bmatrix}$$

For the three equations, the relevant matrices are $A1 = \begin{matrix} -1 \\ -2 \end{matrix}$, $A2 = \begin{matrix} 1 \\ -2 \end{matrix}$ and $A3 = \begin{matrix} -1 & -2 \\ -1 & -1 \end{matrix}$

In this case, the determinants of A1 and A2 are undetermined, whereas the determinant of A3 is -1. So, two out of three equations fail the rank condition and it is therefore safe to say that the model is unidentified.

same combination of cells, in which case the total number of unknown variables is lower than 3821. However, this does not change the inevitable conclusion that the model describing the unknown household-transformations is underidentified. This means that there is no algebraic solution, resulting in the consistent household-transformations.

So, where do we go from here? As it is obvious that there is no algebraic solution to the problem, we could try to find a numeric solution. The only way in which a solution for a certain future year could be found, would be by iteratively adjusting a consistent transformation-variable to reflect the distribution of gender, LIPRO and age-group as much as possible. However, not only would this be extremely CPU-consuming, most likely, but there would also be no guarantee that a solution would eventually be reached. Moreover, it is very well possible that the numerical solution for a certain year would be totally different from the numerical solution of another year, thereby making the simulation results highly unstable over time.

For the above reasons, a ‘second-best solution’ was tried out: the **household-average** growth - or transformation rate of the age-specific growth rates of the indices clearly meets the demand of consistency Let us turn back to the first stylised example to explain what is happening. Remember that the indices are 1.5, 1 and 2 for individuals of 20, 30 and 60 years of age, respectively. Based on these individual transformations , the household-average transformations are

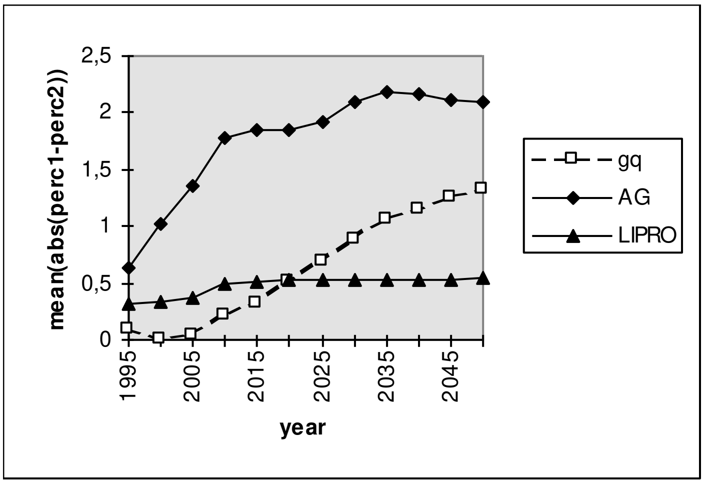
Household 1: (1.5+2)/2=1.75
Household 2: (1.5+1)/2=1.25

So the household-average growth rates of the index are 1.75 and 1.25. for household 1 and 2, respectively. What happens if we transform the population at t with these growth rates?

	t	t+n	true numbers t+n
individuals(age=20)	2	1*1.75+1*1.25=3	3
individuals(age=30)	1	1*1.25=1.25	1
individuals(age=60)	1	1*1.75=1.75	2

As a result of the household-transformation, the simulated sizes of the various age categories (column 2) is no longer equal to the sizes actual of the age categories based on the individual transformations (column 3), which is exactly what we expected, based on the earlier explanations. The question is not if there is such a disturbance of using the household-transformations on the proportional sizes of sample categories (that we know already), but merely what the magnitude of these disturbances is. This will be shown by subsequently subdividing the population into various categories. These categories are gender (male, female), age category (younger than 20, between 20 and 64 and from 65 onward). Denote *PERC1* the proportional size of a certain category given individual transformation. This the 'correct' proportion. Denote *PERC2* the proportional size of the same category, but given the household-average transformation. For the three cases (gender gq, age AG and household-type LIPRO), the average absolute difference between the proportions (i.e. the percentage-point differences) are depicted in the following figure:

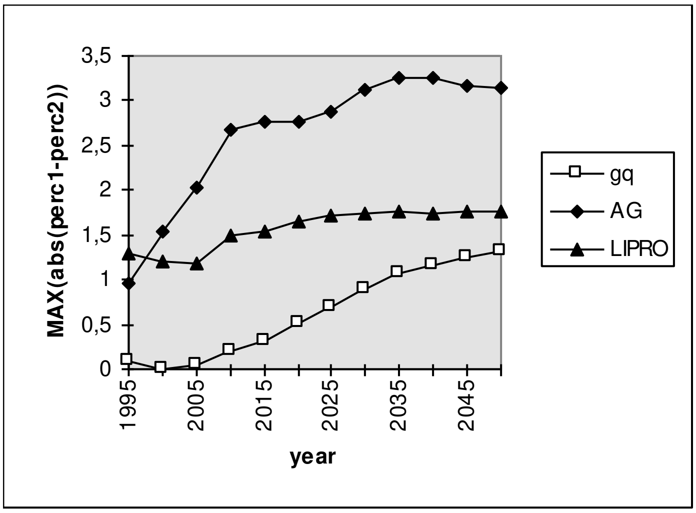
Figure 13: percentage-point differences between individual- and household-average reweighting of the sample.



So, for a certain year, the value for gq (gender) is the average value of the proportional absolute difference for males and females, where this difference is simply the proportion under the assumption of the individual transformation minus the proportion under the assumption of the household-average transformation.

Knowing the average absolute difference is not enough; we might also want to know the maximum value underlying this average difference. These maximum absolute differences is the maximal bias which a subcategory can encounter if the household-average transformation is used instead of the individual transformation.

Figure 14: maximal absolute percentage-point difference between individual and household-average reweighting.



This figure shows that the maximal bias is less than 3.5%-point.

So, the choice we have boils down to the following: either we use the individual transformations with the rather fundamental problem that the figures are based on a non-existing notion of household, or we use the household-average transformations and we accept the biases. Unfortunately, there is no clear-cut answer as it is a arbitrary choice, but it is straightforward that we will have to get an idea on the relative importance of this bias i.e. the size of the bias, caused by the household-average transformation, as compared to the size of the individual transformation. Denote *PERC1992* the proportional size of a certain category in the original, untransformed dataset of 1992. Then the relative importance of the (individual) transformation can be written as ABS(*PERC1-PERC1992*)/*PERC1992.* The relative importance of the disturbance as a result of using the household-average weights is ABS(*PERC2-PERC1*)/*PERC1*. Note that the nominator of this variable formed the basis for the above figures. Then, for all future years, the average and maximum values of both variables can be put in a graph. The sample was then subdivided to age (or one-year age groups) and, next, to household-category LIPRO[24]. For the age-categories and the future years, the figures are:

Figure 15: average percentage bias to age groups.

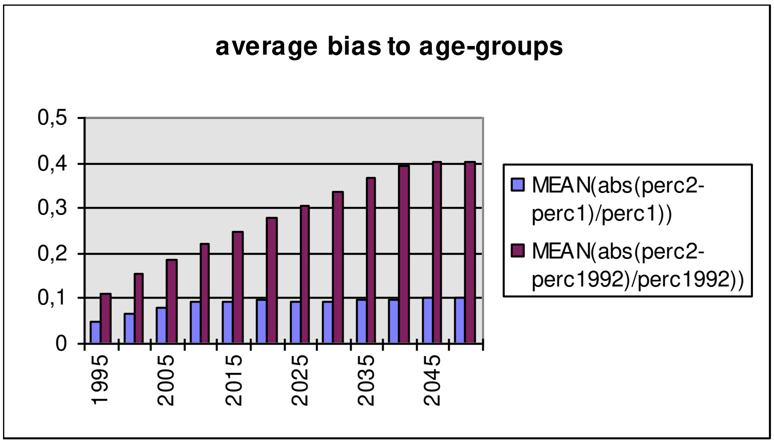
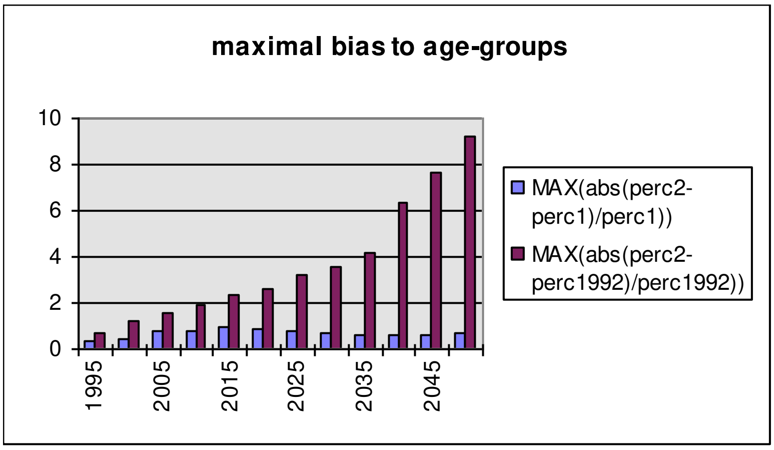


Figure 16: maximal bias to age groups.



The worrying picture sketched by figure 16 is fortunately countered by the preceding figure 15. The maximal bias may not be unimportant, the average bias shows that this maximal bias is found in proportionally small age categories.

Next, the sample is subdivided to household-categories. For the future years, the results are:

[24] The classification to gender was omitted, since it can be expected that the reweighting of the sample has only a very limited effect on the proportional size of these categories. This makes comparison with the bias-effect less relevant

Figure 17: average bias to LIPRO-categories.

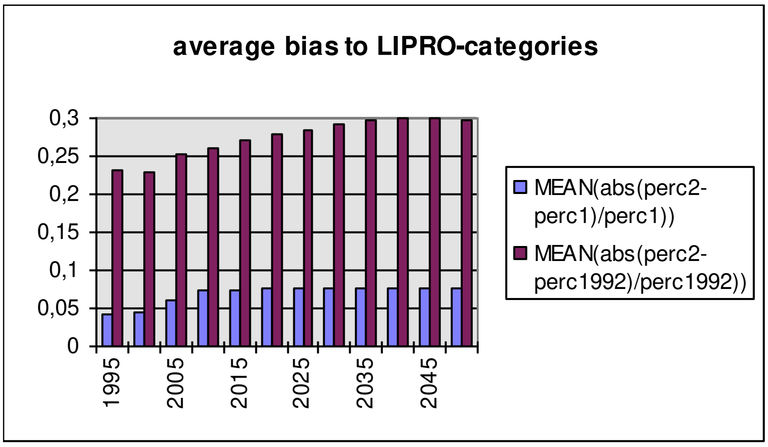
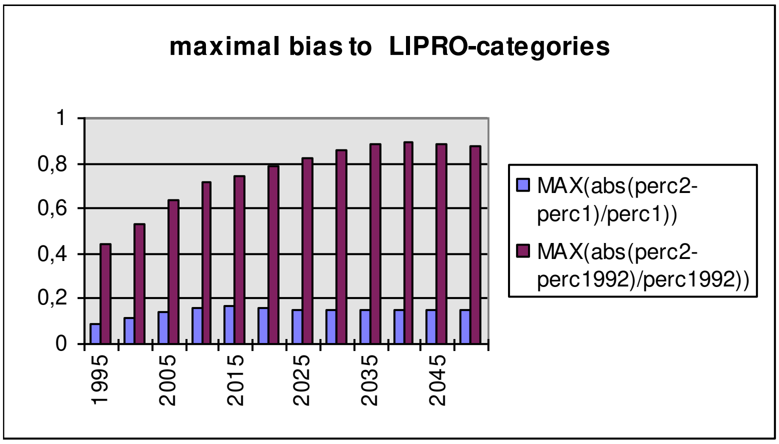


Figure 18: maximal bias to LIPRO-categories.

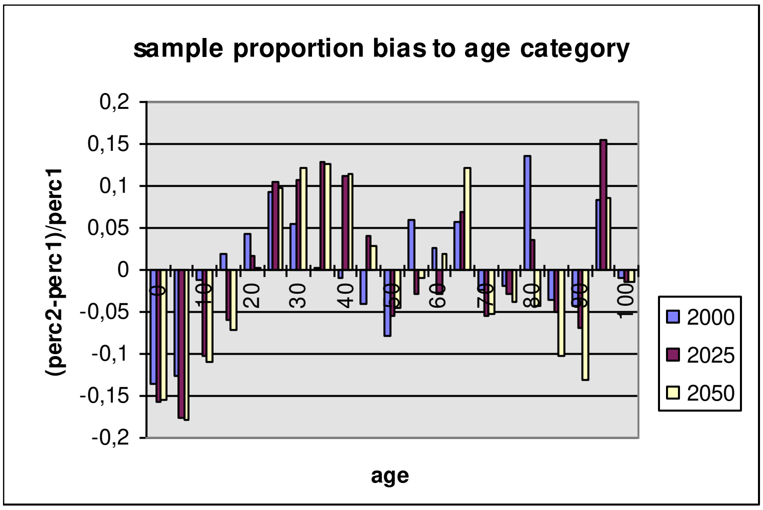


Using the above figures, the conclusion is that the bias is quite small as compared to the relative size of the 'individual weighting-effect'. It is remarkable, by the way, that this bias seems to be increasing the most in the beginning of the simulation period. After that, it reaches a certain maximum and (in the case of the LIPRO-classification) even becomes somewhat lower. In general, the decision to consider the bias as being limited and therefore to accept the household-average transformations as a point of departure for further analyses, seems justified.

Before ending this paragraph, let us briefly take a closer look to the bias caused by using the consistent household-average transformations instead of the inconsistent individual-transformations. In the above figures, the absolute value of this bias was taken, since the goal was to decide whether or not this bias was considered small enough to be acceptable or not. Next, let us consider the actual value of this bias, so that we can see what categories are overestimated and what categories are underestimated, even if these over- and underestimations are limited. For three years

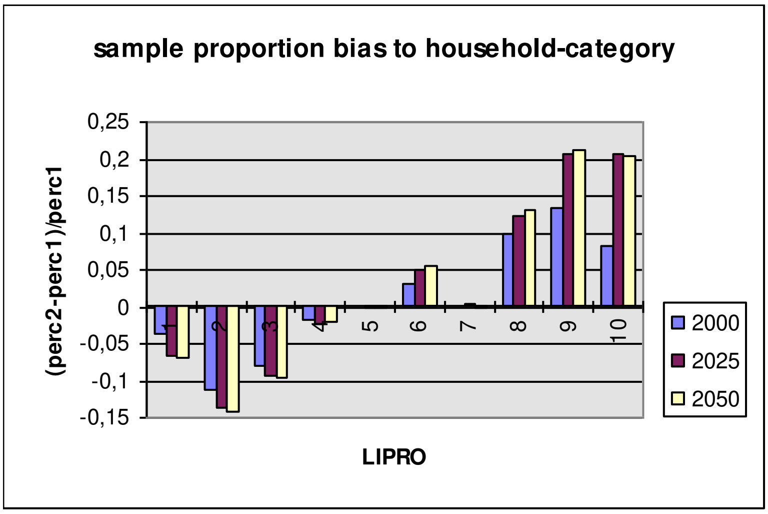
(2000, 2025 and 2050) and the same classifications as used above, the actual values of the bias are given in the next two figures.

Figure 19: sample proportion bias to age categories.



First of all, the direction and size of the bias is more or less the same for the three years under consideration. This holds as well for the household-categorical-bias as presented in the figure below. Secondly, as a result of using the consistent household-average transformations, the middle-aged individuals are somewhat overestimated, at the expense of the youngest individuals. A possible explanation for this is that middle-aged individuals (from 20 to 45 years old) are often in the same household as children under 20. So, if the one category becomes overestimated as a result of using the household-average transitions, the other category must by definition become underestimated. But this only explains why there could be an inverse relation between the young and middle-aged individuals, in terms of over- or underestimation. However, why isn't it so that the young individuals are overestimated (and the middle-aged individuals underestimated)? A possible answer can be found in table 3: over the whole range of future years, the growth rate of the proportion of young is higher (less negative) than that of the middle aged individuals. So, relative to each other, the proportion of the young individuals increases whereas the proportion of the middle-aged individuals decrease. Now if the assumption that individuals of these two groups are likely to be found in the same households, the use of household-average transitions will result in an underestimation of the young individuals and an overestimation of individuals belonging to middle age-groups. For older individuals, the picture is less clear-cut: some age-categories (50, 70 and 85-90) and underestimated whereas others (65, 80 and 95) are overestimated. Moreover, for the later future years, the underestimations become larger at the expense of the overestimations. This weak intertemporal movement is most likely caused by old-aged individuals living in the same household as young and middle-aged individuals. Figure 4 shows clearly that the individually-transformed proportions of retirees grow strongly, whereas the proportions of young- and middle-aged decrease. So, if household-average transformations are used, the retirees should be gradually more underestimated and less overestimated as time goes by.

Figure 20: sample proportion bias to household-category.



In the case of the household-classification (LIPRO) is can be seen that using the consistent household-average transformation results in a bias which underestimates the children (which is closely related to the underestimation of the young individuals, as shown in the above figure), whereas LIPRO-categories 8, 9 and 10 (Cohabiting individuals with children, single parents and other cohabiting individuals) are somewhat overestimated. Fortunately, the most important categories (singles, married individuals with and without children) are nearly unbiased.

2.5. Conclusion.

In this second chapter, the main characteristics of the static microsimulation model STATION have been presented and discussed. The model combines both fundamental techniques involved in the static ageing of a dataset: reweighting and uprating. Using exogenous demographic projections, the model reweights the 1992-dataset, more specifically the age- and household-type distribution. This describes a world where the macroeconomic situation is 'frozen' and where only the demographic situation changes. Of course, this scenario is just an illustration, since it is highly unrealistic. To develop more realistic scenarios and establish a link with the time-series model of the Belgian Federal Planning Bureau, the technique of uprating is used to implement the assumptions which the Belgian Federal Planning Bureau used to build their model of the future costs of the Belgian pension system, but for which they did not (sufficiently) explore the distributional effects on income. These assumptions give way to three scenarios. As opposed to the benchmark-scenario, all of these scenarios share the assumption that wages increase by 2.25% per year, in real terms. In the **first scenario** or the scenario of no policy change, it is assumed that pension benefits are not linked to this development of wages. This scenario resembles the most to the current situation where pension benefits are linked to the rate of inflation but not to the wage-rate. The **second scenario** implements a partial linkage between pensions and wages. Finally, the **third scenario** introduces the so-called wage-barrier in the analysis, and assumes an annual increase of this wage-barrier of a certain fraction of that of the annual increase of wages. One should however keep in mind that these simulation variants all describe the situation, which starts in 1993. The past situation is implicitly reflected in the 1992-dataset and can therefore not be changed, nor 'wiped out'.

After the presentation and discussion of these scenario's, this second chapter continued with describing some ad-hoc problems involved in static microsimulation, and discussed the solutions found to overcome these problems. This study now continues with considering the simulation results.

Chapter 3. Simulation results.

3.1. Introduction.

In the first chapter, the problem was defined: it was argued that the ageing of the Belgian population could have profound effects on the pension income. Moreover, it was argued that 'conventional' models did only show the development of pension income over time, while ignoring the distributional effect of ageing on pension income. In the second chapter, the static microsimulation model STATION was presented and thoroughly discussed. In this third chapter, let us turn to the presentation and discussion of the simulation results. But let us take a small side-step and dwell upon the notion of 'welfare' for a moment. The reason why this is, is because the simulation results presented below, are in some way related to 'welfare'. The reader who is interested in a theoretical discussion of the matter, is referred to the book by Barry (Barry, 1990) whereas the article by Slesnick (Slesnick, 1998) gives an overview of the empirical aspects of welfare-measurement. The vast majority of studies dedicating to the study of welfare of households, analyses the effect of prices and incomes on consumer's surplus. However, due to a large number of theoretical and practical problems of which the problem of aggregation over households is just one, attention shifted towards the analysis of incomes and expenditures. Expenditures are intuitively more closely associated to 'welfare' than incomes, but incomes require less assumptions concerning prices and such, and are within the direct 'policy-reach' of the government, namely via the tax and transfer system. So, income is the point of departure of the analysis: household income, to be exact. The reason why household-income is chosen over individual income is that the latter is expected to be more related to consumption- expenditures of individuals than the former. For example, numerous individuals who do not work and earn a labour-income (children, for instance, or housewives) do consume out of the income of income-earners in the same household, and therefore gain utility. Moreover, tax- and benefit regimes are often more household- than individually-based. However, this poses a certain problem, since we can imagine that welfare differences emerge between households of different sizes but the same household income. Put differently, if two households earn the same income, we can imagine that the level of welfare differs if the one household consists of, say, seven individuals whereas the other is an one-individual household. Consequently, simply comparing household incomes is not good enough. So-called 'equivalence-scales' must be applied to the household-income, in order to increase comparability. Slesnick (1998, p. 2147) defines an equivalence scale as "*the expenditure, relative to a reference household, necessary to attain* [a certain, GD] *utility level*". Numerous equivalence scales relating subjective welfare to household-income have been developed. However, in this study, a very simple and (therefore) commonly adopted equivalence scale, the so-called EU-scale, is used. This scale gives the head of the household a value of 1, other adults a value of .5 and children a value of .3. So, the equivalent household-income of a household consisting of a father, a mother and two children is derived by dividing the household-income by 1+.5+.6=2.1. From now on, unless stated otherwise, when income is discussed, I mean equivalent household income.

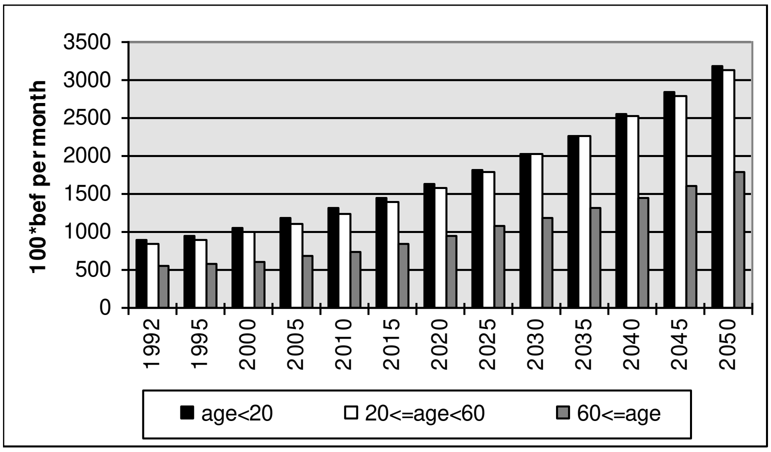
This chapter starts by showing the development of average total household income of various age-groups. First, the average income of those older than 60 is compared with the income of younger age groups (younger than 20 and between 20 and 60). This way, the effect of the uprating process as well as the macroeconomic budget equation become highlighted. Next, we will concentrate on the group of retirees: the development of average income of this category -subdivided into age-groups of course, will be discussed.

Taking the average income as the notion of welfare is, of course, a rather rough analysis, assuming a utilitarian welfare function and a constant marginal utility of income (Slesnick, 1998, p. 2109). The distribution of income should be taken into account as well, since "*the concept of vertical equity is often identified with the notion that, all other things equal, more egalitarian distributions are preferred to those that are more dispersed*". This concept of vertical equity will be included in various ways. First of all, the intertemporal distribution of relative poverty rates will be presented and discussed. Here, the percentage of 'poor' will be derived, defining one to be poor if he is in a household whose income is below half the average household income. Implicitly, this incorporates vertical equity, since "*a decrease in the dispersion of incomes would tend to lower the number of individuals classified as poor*" (Blackburn, 1998, p. 453). Next, the inequality of income as such will be presented. More specifically, I will consider the development of the Theilcoeficient over time and given the four simulation variants under consideration, both for the population as a whole and for subgroups.

Up to then, only partial measures of welfare will have been introduced. Either we will have discussed the average income or the distribution of the average income around the mean. Of course, this is a simplification of the facts, since there most likely will be some sort of trade-off between level and inequality. In the second part of this simulation results-section,. Based on a welfare function developed by Sen, which explicitly introduces this trade-off between income-level and inequality, Kakwani (Kakwani, 1986, p. 212-213) developed additive welfare functions for subgroups, using a subjective notion of ‘envy’. This section will end by presenting and discussing the intertemporal development of this last variable for retirees and non-retirees.

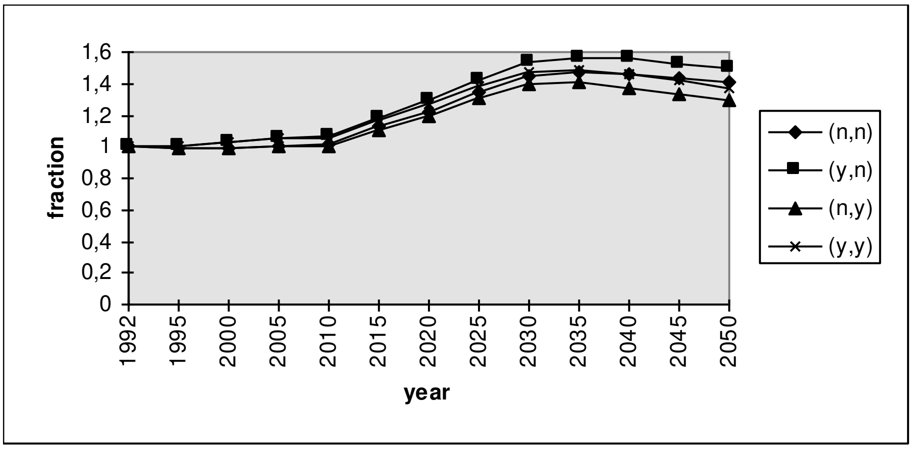
As said, let us start by comparing the development of total equivalent household-income, as averaged over individuals, over time. First of all, let us consider the development of income if there is no macroeconomic identity incorporating the PAYG-element of pensions. By doing so, we get a chance to see the impact of the above-discussed uprating-equation on income.

Figure 21: development of average income, no indexation, no wage-barrier, no macroeconomic PAYG-identity.



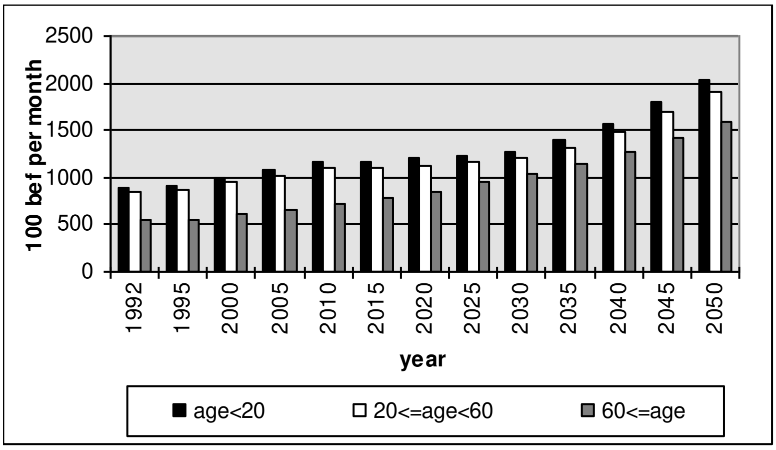
Remember that the assumption is that wages increase by 2.25% per year, whereas the incomes of the retirees are not adjusted (in real terms, that is). This is why the income of the retirees gradually lags behind the incomes of the non-retirees. Nevertheless, it is clear that the average income of the retirees increases as well, since the average pension base of new pensioners increases over time. Next, what happens when the PAYG-budget identity is included? Before considering the effect on the average income, let us consider the effect of the income of the retirees on the income of the non-retirees.

Figure 22: effect of the income of the retirees on the income of the non-retirees.



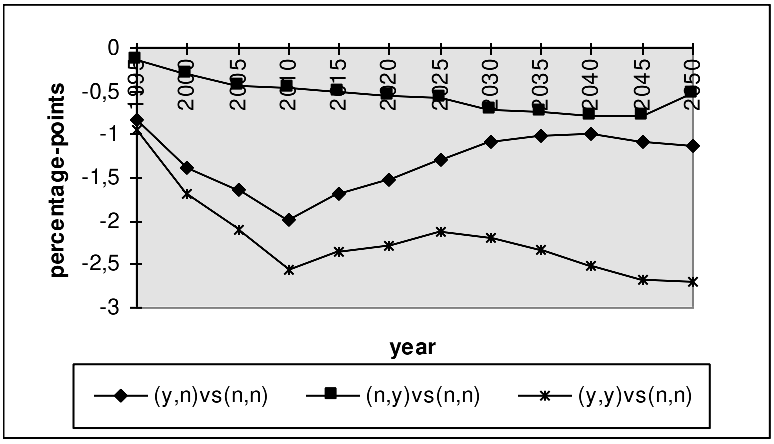
In order to let the incomes of the non-retirees take the contributions necessary to provide the increasing total income-mass of the retirees i.e. the inclusion of the PAYG-budget equation, the incomes of the non-retirees must be divided by the above ratio. As a result of both demographic ageing and uprating and given the simulation variant where there is no indexation and no wage-barrier, the income-mass of the retirees as a fraction of the income-mass of the non-retirees will increase with at most 48% in 2035. Taking this into account, the development of the average incomes of the three age-categories under consideration is shown in figure 23.

Figure 23: average total income (no index, no wage-barrier)



Up to about 2010, the same pattern as in figure 20 emerges. But then ageing really kicks in, causing the average incomes of the non-retirees to decrease relative to its development in figure 20. The incomes of the retirees and non-retirees to some extent converge up to 2035. After that, the situation normalises in the sense that the incomes slowly diverge again. In other words, even if the incomes of the non-retirees growth by 2.25% per year whereas the incomes of the non-retirees remain constant (in real terms), then the diverging effect this has on the incomes of the retirees and non-retirees is between 2010 and 2035 partially neutralised by the negative effect of the increasing costs of ageing on the incomes of the non-retirees. Next, let us consider what happens if the pension benefits of the pensioners become indexed (variant 2), if the wage which forms the basis of the future pension benefit becomes limited (variant 3) and both (variant 4). It can be seen from figure 22 that variant 2 results in an increase of the cost of the pension-system for the non-retirees (up to 8.6%-points in 2035), whereas the imposition of the wage-barrier results in a relative decrease of this cost (up to 7.3%-points). As a result of the joint implementation of both measures, the cost of pensions will more or less remain the same (a maximal increase of .86% in 2035) relative to the situation where there is no linkage and no wage-barrier. These policy measures clearly show the trade-off between the income of the retirees and non-retirees. Instead of simply presenting the average incomes for the three age-categories in different figures, which would all be very much alike and therefore quite uninformative, let us evaluate the effect of these measures in terms of *Pareto-optimality*. A situation is said to be Pareto-optimal if it is not possible to increase the utility of one individual or group of individuals without decreasing the utility of another individual or group of individuals. This definition has been made operational using compensation mechanisms. So, we have to consider whether or not the gainers from a certain measure could compensate the losers in terms of utility, where utility is again represented by average income. Let us ignore the youngest age-category, as their 'behaviour' in terms of changing average income as a result of the policy measures is about the same as that of the category aged between 20 and 64. Then, the percentage-point difference between the growth rate of the average income of the winners and that of the losers of each simulation variant, is given in figure 24.

Figure 24: Pareto-efficiency of policy measures



As figure 24 clearly shows, neither the partial or combined implementation of the (partial) linkage of the pension benefit and the (partially-linked) wage-barrier is a Pareto-improvement[25]. But what does this tell us? Not very much, unfortunately: only that the average income of one category decreases more than that the average income of the other category increases. It does not say anything about who the winners and losers are, nor does it say anything about how rich, respectively poor these categories will be after the implementation of these measures. This is why the joint implementation of both measures results in a higher diversion from the Pareto-optimal situation as compared to the situation where only the pension benefit is linked to the development of wages. At first, this is awkward, since the implementation of the wage-barrier decreases the loss of the non-retirees while simultaneously decreasing the gain of the retirees. This is true, but we are not considering the direction of the losses or gains, but only their *relative magnitude*. In fact, the decrease of the loss of the non-retirees is smaller than the decrease of the gain of the retirees causing the balance to be lower[26]. So, what can we tell about the joint implementation of more than one policy measure, versus the effect of implementing one of these measures? Not much, not to mention 'nothing at all'. To make things worse, if the winners and losers change places 'along the way'[27], what conclusions are there left to draw? This is one of the reasons why this kind of analysis is not very useful. However, when we stick to the partial or one-at-a-time implementation of the linkage, respectively the wage-barrier, and if we assume a constant marginal utility of income while ignoring the effect of the distribution of income on welfare, we can conclude that the losses outweigh the gains.

---

[25] Note that the 'winners' and 'losers' are not always the same, of course: if the wage-barrier is implemented, the non-retirees are the winners ((n,y)vs(n,n)) whereas the situation is reversed in the case of the partial linkage of the pension benefit ((y,n)vs(n,n)). However, in either cases, the gain of the winner does not compensate the loss of the loser.

[26] Giving a numerical example might clearify this. In the case of only the partial linkage of the pension-income to the course of wages, the 2010-gain for the retirees is 3,25%, whereas the loss for the non-retirees is about 5,25%. The balance is therefore -2%-points. When both the partial linkage and the wage-barrier are implemented, the loss of the non-retirees decreases to 4,25% whereas the gain of the retirees decreases to 1,7% -a considerably higher number. So, the balance becomes about -2,55%.

[27] which is in fact the case in the case of the joint implementation of the partial linkage and the wage-barrier: at first the eldely are the winners, as the effect of the wage-barrier becomes stronger only gradually. But after 2030, the situation is reversed. This will be shown later, in figure 27.

Next, let us consider the same average household data, but then with another categorisation of age. In figure 25, the development of income of the retirees is highlighted.

Figure 25: average total income (no link, no wage-barrier)

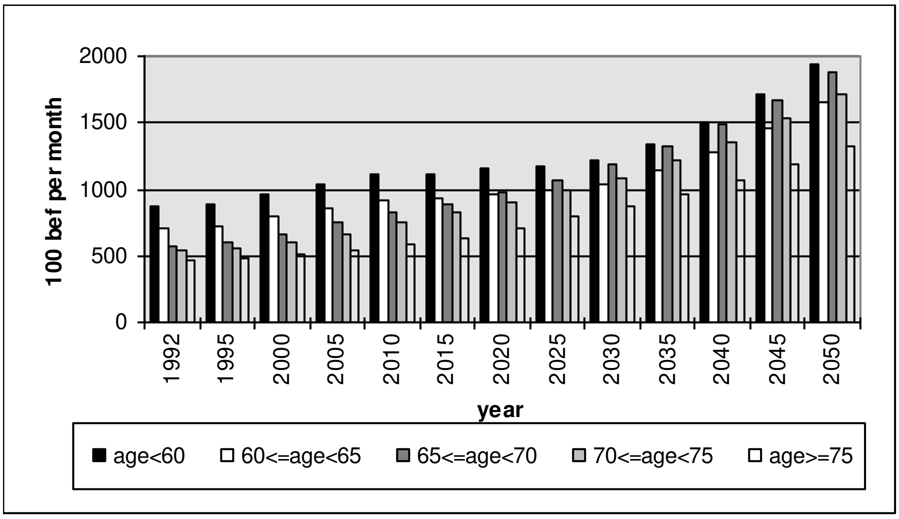
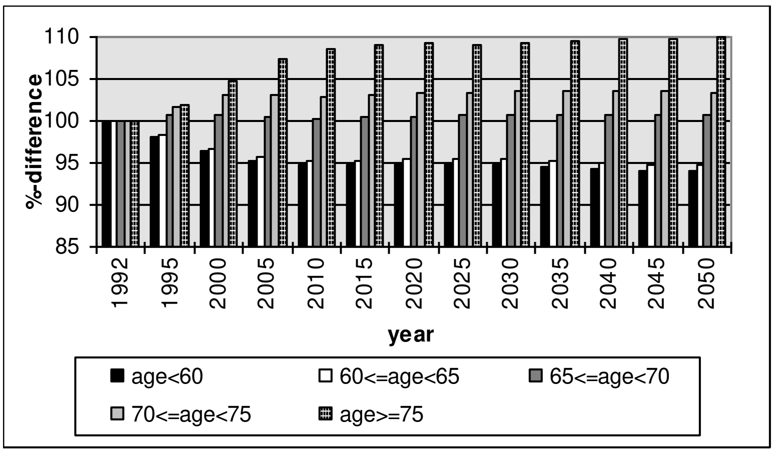


Figure 25 again shows the pattern which emerged in figure 23 as well. At first, the incomes of the non-retired diverge from that of the retired. Then, from about 2010 on, the growth of the incomes of the non-retirees slows down, resulting in a convergence of the income of the non-retirees and the retirees. However, the average equivalent household-income of the young retirees becomes higher than that of the old non-retirees from 2020 on. This turns out to be persistent and quite strong since the diverging process which sets in from 2035, and which was shown in figure 23 as well, is not strong enough to restore the old situation. Why this is so, becomes immediately clear by noticing that we are dealing with *net* incomes here. The assumption of a 2.25% growth rate was based on a expected growth rate of the economy as a whole and therefore applied on gross incomes. If the value of α changes for some reason, the net income of the non-retirees will decrease relative to the gross income. But as retirees do not pay pension-contributions (only sickness- and disability benefit, which is assumed to be constant), their income is not adjusted downwards. As a result, it is possible that the net income of somebody who retires, increases since he or she does no longer have to contribute to the pension system. The reason why the average income decreases as the age-category of the head increases, is closely related to the upgrading processes as described in the third chapter. For any future (or current) year, the older a pension beneficiary is, the more his or hers pension benefit has been following the course of wages, relative to older age-categories.

Next, let us consider what happens if we apply the three scenarios as described in chapter 3. Remember that the base-variant was the situation where ongoing pension benefits are not linked to the wage-rate, whereas wages increased by 2.25% per year. Moreover, the wage-barrier is assumed to increase at the same speed as the wage-rate which in this case is equivalent to saying that the wage-barrier is linked to the rate of inflation, or that there is no wage-barrier at all (since the model is in real terms). As shown in table 1, if we describe each variant with two letters, this base-variant is described by **(n,n)**: **no** linkage to the wage-rate and **no** wage-barrier. In the second variant, we follow the Federal Planning Bureau in assuming a partial linkage between the course of wages and the course of the pension benefit. In fact, we assume that ongoing pension benefits (that is, pension benefits of individuals who are retired) increase by 1% per year. Again, there is no wage-barrier. To summarise, this second variant is described by **(y,n)**. For the third variant, we assume no linkage -as in the first variant- but this time with an assumed annual increase of the wage-barrier of 1% per year **(n,y).** Lastly, it is both assumed that pension benefits are linked to the wage-rate, and that the wage-barrier increases by 1% per year **(y,y)**. Having recapitulated the simulation variants, let us consider the effect on the average pension benefit, expressed as a fraction of the (n,n)-pension benefits. First of all, what is the effect of linking the income of the retirees to the development of (gross) wages of the non-retirees? This is shown in figure 26 where the relative figures are presented.

Figure 26: the effect of a partial linkage on average total income.



Clearly, introducing the linkage of the incomes of the retirees to the development of gross wages is the most beneficiary to the oldest category of retirees. When discussing the simulation results of the (n,n)-scenario, we argued that the lower income of the oldest category of beneficiaries was caused by the fact that the older a beneficiary is, the more his or her pension benefit lags behind the current level of wages. Partially linking the pension income of the retirees to the wage-rate prevents this lagging behind and it is therefore clear that the age-category for which this lagging was the strongest -the oldest age-category- benefits the most.

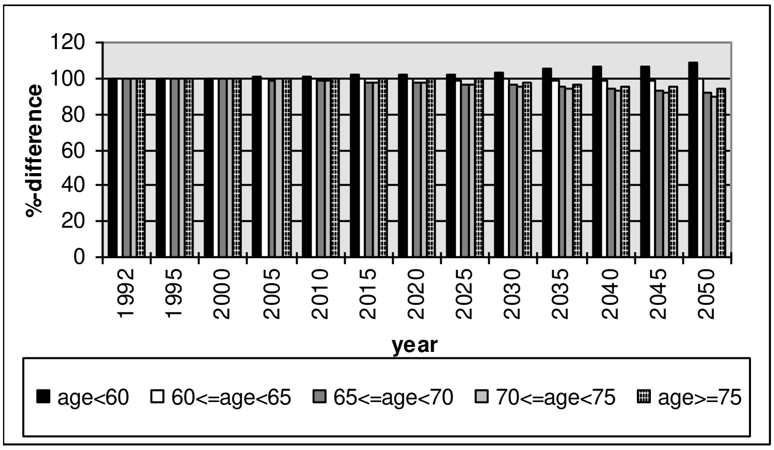
Another remarkable thing is that the effect of this partial linkage is the strongest for the earlier simulation years. This effect is *not* caused by the uprating mechanism itself. To see why this is so, let us rewrite the uprating equation presented earlier. As before, denote *y* the future year and *a* the age of the individual in the future year *y*. Moreover, denote *x*=*y*-1992 and *age*=*a*-*x*. Then the upgrading equation in the case that the ongoing pensions are increased by 1.% per year can be rewritten as $(1.0225)^{x}$ for $a<65$ and $(1.0225)^{\max[0,65-age]} \times (1.01)^{\max[0,x-\max[0,65-age]]}$ for $a\geq 65$. This is the second- or (y,n)-variant in paragraph 2.3. If, on the other hand, there is no partial indexation of the pension benefit to the course of wages, the upgrading equation becomes $(1.0225)^{x}$ for $a<65$ and $(1.0225)^{\max[0,65-age]} \times (1.00)^{\max[0,x-\max[0,65-age]]}$ for $a\geq 65$. Denote this the (n,n)- or the benchmark-variant, as presented in paragraph 2.3. The above figure shows the relative effect of partially indexing the pension benefit for different age-categories. This can be written as $R=(y,n)/(n,n)$. Write this out to see that $R=1$ for $a<65$. For $a\geq 65$, R can be written as $R=((1.01)^{\max[0,x-\max[0,65-age]]})/((1.00)^{\max[0,x-\max[0,65-age]]})$. Rewrite to $R=((1.01)^{\max[0,x-\max[0,65-age]]})/(1.00)$ or $((1.01)^{\max[0,x-\max[0,65-age]]})$. So, the marginal effect of x on *R* is log(1.01)=.004. At first sight, this is somewhat awkward, since a constant marginal effect implies that the income of the retirees in the case of a partial linkage should increase at a constant speed, relative to the pension income in the case of no linkage, ceteris paribus. However, this ceteris paribus clause *is not met*, since the microsimulation model causes the population to change with respect to the variable *age*. Especially in the first simulation years, the relative weight of older individuals increases as a result of ageing. As said before, the older the individual is, the more he or she gains from indexing. So, *R* is higher for older individuals. If the relative proportion of older individuals increases *within an age-category* as a result of ageing, the 'average *R*' for that age category increases, which is what we see in the above figure.

As said earlier in this study, the speed of ageing is the highest in the early simulation years: from about 2020 onwards, the relative proportion of the oldest individuals will decrease again (especially in the oldest age-category). As a result, the average *R* will decrease again, meaning that the effect of the uprating process will stabilise. This also explains why the relative effect of the partial linkage stabilises earlier for the younger age-categories of retirees. For the category between 65 and 70, the effect stabilises quite early in the simulation period, namely around 1995. For the other age-categories, this turning points respectively are 2000 and 2015. This is because there is not only a shift within the age-categories, but also between the age-categories: if we abstract from the static nature of the model for a moment, we could say that individuals transit from an age-category to a higher age-category. So, ultimately, the effect of the linkage shifts to the oldest age-category.

We have not yet discussed the effect of the partial linkage on the incomes of the non-retirees: the fact that the net incomes of the non-retirees decrease as a result of the introduction of the partial linkage, is not unexpected, of course. What is interesting, though, is that the above figure indirectly supports the notion that this policy measure is not Pareto-optimal. Of course, the gain of the oldest-age category is larger than the loss of the non-retirees, but this category of retirees is small. If we take into account that the vast majority of retirees are younger than 75, it immediately becomes clear that the gains do not outweigh the losses.

Next, we consider the (n,y)-variant, where there is no linkage, but where a wage-barrier is introduced instead. The simulation results are shown in figure 27.

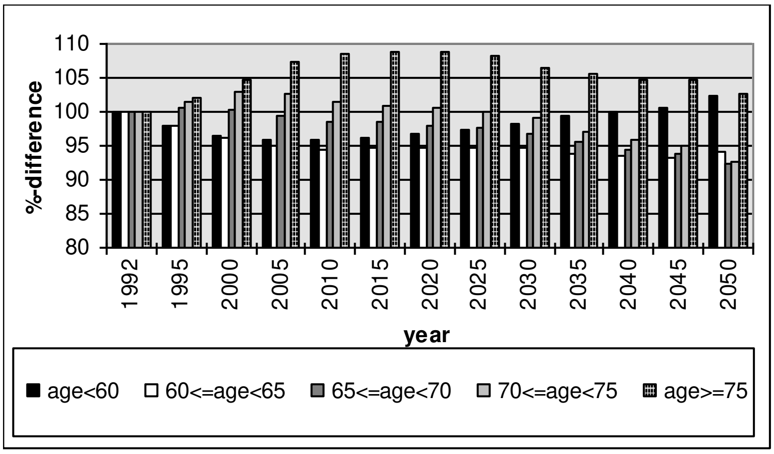
Figure 27: the effect of a wage-barrier on average total income.



The fact that the incomes of the retirees are lower than in the case of the (n,n)-scenario is a direct result of the introduction of the wage-barrier, which limits the pension benefits of the highest-earners to 1% per year, whereas the (gross) incomes of the non-retirees increase with 2.25% per year. So, the effect of introducing the wage-barrier becomes stronger over time. The fact that the incomes of the retirees gradually decrease, implies of course that the incomes of the non-retirees increase as they doe not have to contribute as much to the income of the retirees as they should have in the case of no wage-barrier. Note that the negative effect is the strongest for the youngest retirees. The reason lies in the fact that the gross wages of the non-retirees decrease by 2.25% per year, whereas the wage-barrier increases by 1% per year. This means that the wage-barrier decreases in relative terms. So, the further in the future we get, the more individuals will see the wage-barrier influence their pension benefit. This means that, given a certain future point in time, a larger proportion of the young retirees will have encountered the wage-barrier than compared to the older retirees.

Finally, the development of income of the various age-categories in the case of the simultaneous introduction of the partial linkage between wages and pensions, the (y,y)-variant, is given in figure 28 below.

Figure 28: the effect of a partial linkage and a wage-barrier on average total income.



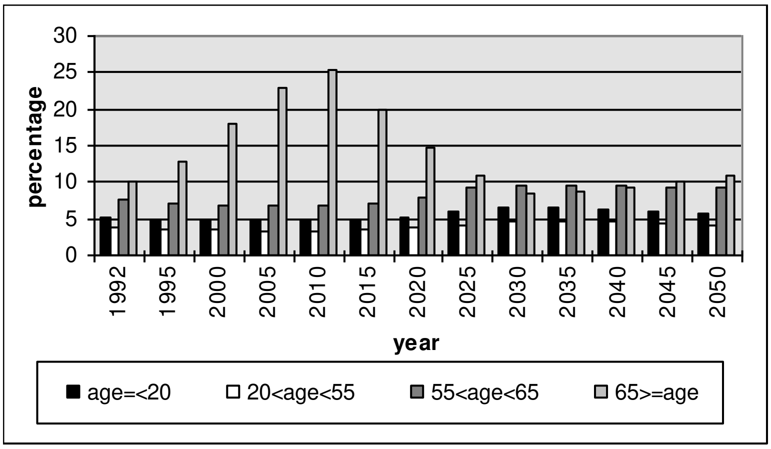
The implementation of the wage-barrier causes a relative increase of the income of the retirees, whereas the wage-barrier results in a decrease of their incomes. Of course, via the macroeconomic PAYG-equation, the income-development of the non-retirees is the reverse of that of the retirees. The joint implementation of both measures therefore has an ambiguous effect (see footnote 26). Here, an interesting pattern emerges: the effect of the wage-barrier on the income of the retirees emerges only slowly, as it takes effect when today's non-retirees reach the retirement age. The partial linkage of the income of the retirees to the incomes of the non-retirees immediately takes effect. So, in the short run, the same pattern as shown in figure 26 emerges: the incomes of the retirees increase, where the oldest age-categories gain the most, whereas the incomes of the non-retirees decrease. However, from about 2015 on, the situation changes. Ageing has lost its momentum, and the effect of the wage-barrier has gradually become stronger. As a result, the incomes of the retirees start to decrease again. The incomes of the non-retirees eventually even increase relative to their 1992-level.

When discussing the effect of the partial linkage of the incomes of the retirees to the incomes of the non-retirees, it was argued that the reason why the positive effect of this linkage on the average income of the age-category reaches its maximum (i.e. the year from which it remains more or less constant) later in the simulation period if the age-category is older: this was because ageing causes not only the oldest-age category to gain importance relative to the other age-categories, but also by the fact that the average age increases within each age-category. Moreover, when discussing the partial effect of the wage-barrier, we concluded that the negative effect of this barrier is the strongest for the youngest category of retirees. The combination of both measures therefore results in a relative shift of income from the non-retirees and the young retirees to the old retirees. For the last 15 years of the simulation period, the wage-barrier will have gained so much strength, that the contributions which the non-retirees make will have decreased relative to their 1992-level. From this year on, the non-retirees and oldest category of retirees will gain at the expense of the other two categories of retirees. Note again that the relative differences with the 1992-level (i.e. 100) roughly supports the conclusion that these measures are not Pareto-efficient, as the losses of the losers do outweigh the gains of the winners.

Up to now, we have implicitly assumed that the welfare of (subcategories of) retirees was reflected by the development of average income. Of course, as was explained in the introduction of this chapter, this is a simplification. We will now include the distribution of income in our analysis. At first, this will be done only indirectly, since the subjective poverty rates are to be considered. This is one of the things for which the microsimulation model was developed in the first place. As said in the introduction of this chapter, a rather crude but widely-accepted objective definition of poverty is used: a household is considered ‘poor’ if its equivalent income is lower than 50% of the average equivalent income. So, households are poor relative to other households in the dataset. One should keep this in mind when considering the effects of potential policy measures on poverty. As Heinrich points out, whether poverty is defined in the relative or absolute sense, has important policy implications, since "*absolute poverty is eliminated by making everybody better off i.e. by shifting the income distribution upwards. Relative poverty, on the other hand, is eliminated by redistributing income from the rich to the poor*" (Heinrich, 1998, p.6). As we are primarily interested in the income position of the retirees, we will devote the most attention to this category of individuals in the dataset[28].

Let us start by considering the development of poverty of four age-categories, where the retirees are again opposed to several categories of non-retirees.

Figure 27: poverty rates (no partial linkage, no wage-barrier).



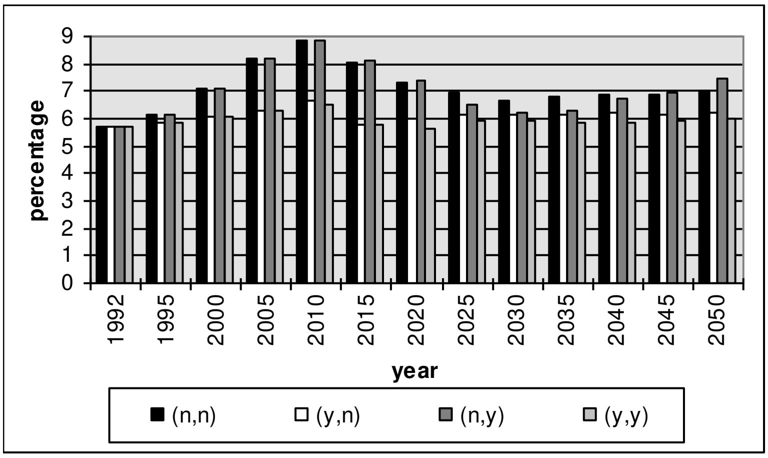
To a certain extend, the development of poverty can be explained by looking back at the figure 25, and roughly comparing the development of income of the non-retirees with that of the retirees. Clearly, as a result of the fact that the incomes of the retirees and the non-retirees diverge over time, the proportion of retirees considered to be poor will increase dramatically up to about 25% in 2010. After that, as explained when discussing figure 25, the cost of ageing cause the incomes of the retirees and non-retirees to converge again, in the sense that the incomes of the non-retirees decrease relative to the incomes of the retirees. Consequently, the proportion of poor retirees decreases again whereas the proportions of poor non-retirees increase. Of course, this

[28] A short note must be devoted to the recent development of combining bootstrapping techniques with poverty analysis. (see, for instance, Heinrich, 1998 and Osberg&Xu, 1997). Poverty analysis is generally based on sample information, which means that there is vulnerability to sample bias and such. One of the solutions to this problem which is gaining popularity due to its generality, its distribution-independence and the technical-simplicity is bootstrapping. However, we will not use this technique in this model. The reason is that de development of poverty for all simulation years is based *on the same sample*, which means that the comparison of the poverty rates are not shaded by sampling errors.

latter development is much less strong as the former, since the category of non-retirees is bigger than that of the retirees, in terms of individuals.

How does poverty changes if the simulation variants are taken into account? Let us start by comparing sample-wide poverty rates given the four simulation variants.

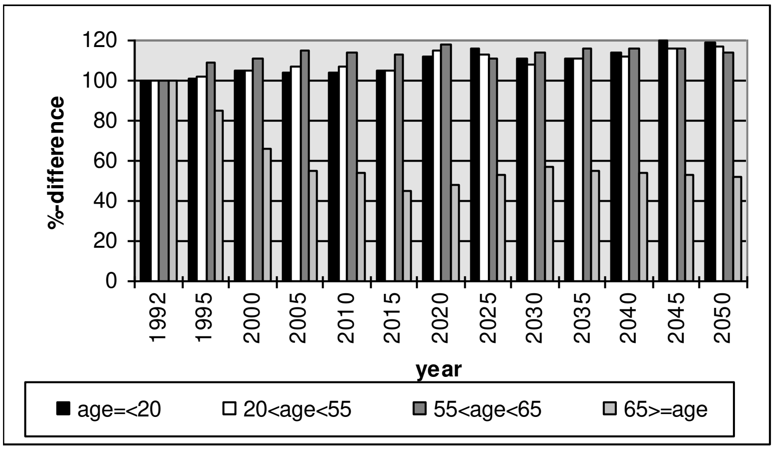
Figure 30: general poverty rates for the simulation variants.



Two major effects immediately become clear in figure 30: first of all, partially linking the income of the retirees to that of the non-retirees strongly reduces the general poverty rate, especially in the first half of the simulation period. This difference is at most 2,25%-point or 39% in 2015. Secondly, the partial implementation of a wage-barrier has only a very small and ignorable effect on poverty. The explanation of the former effect is simple: partially linking the income of the retirees to that of the non-retirees neutralises the divergence-process and therefore reduces poverty. Of course, the implementation of the wage-barrier reduces the income of the retirees, but the incomes which are affected by this measures are those of the richest retirees. Therefore, this measure has only a small effect in the determination of the average income, and therefore on the poverty line. Moreover, this negative effect is partly neutralised by the positive effect of this measure on the incomes of the non-retirees. Of course, one could argue that this is also the case for the (y,n)-variant, where the partial linkage is implemented. In this case, the positive effect on the income of the retirees is opposed by the negative effect on the income of the non-retirees. This is true indeed, but figure 24 shows that this neutralising effect is much weaker than in the case of the (n,y)-variant. So, if this line of reasoning is true, then the age-specific poverty rates should change as a result of implementing the wage-barrier. However, as compared to the changes when the partial linkage is introduced, these changes should more or less cancel each other out. We will consider later if this is so.

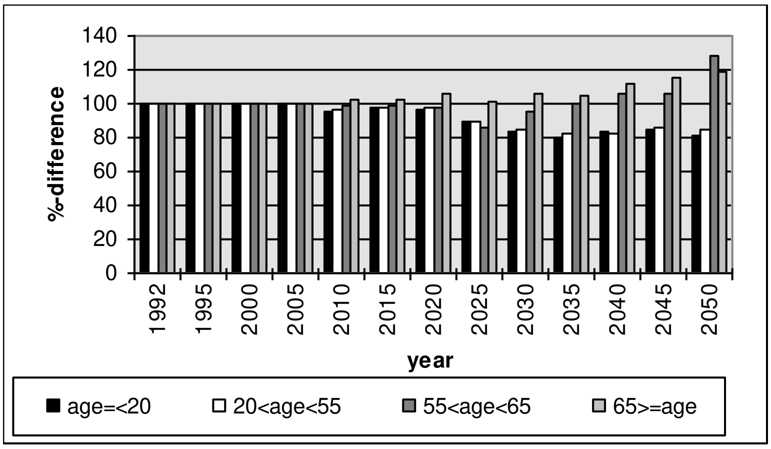
Of course, the next step is considering how the poverty rates of the various age-categories change as a result of the simulation variants. First of all, figure 31 shows the age-specific poverty rates in the case of the (y,n)-variant, i.e. when the incomes of the retirees are partially linked to the incomes of the non-retirees.

Figure 31: the effect of a partial linkage on poverty rates.



As could be expected, the introduction of a partial linkage between the incomes of the retirees and those of the non-retirees results in a sharp decrease of the poverty among the retirees. Eventually, the poverty rate of this category even goes down to 50%-point of its value in the base-variant! This is a combined effect of the relative increase of the income of the retirees and the relative decrease of the income of the non-retirees. As the poverty rate is based on the population-wide average income, we have a zero-sum situation in the sense that a decrease of the poverty of a certain category must be accompanied by an increase of the poverty of one or more other categories. Moreover, the PAYG-equation 5 guarantees that the extra money which is involved in the linkage of the pension income to the course of wages, is paid for by the non-retirees. This is confirmed by the data: the poverty rates for the other categories increase slowly. For the households for which the head is 55 or younger, poverty ends up a bit less than 18%-point higher than in the base-variant at the end of the simulation period. Next, let us consider the effect of implementing the wage-barrier on poverty rates. This is shown in figure 32.

Figure 32: the effect of the wage-barrier on poverty rates.



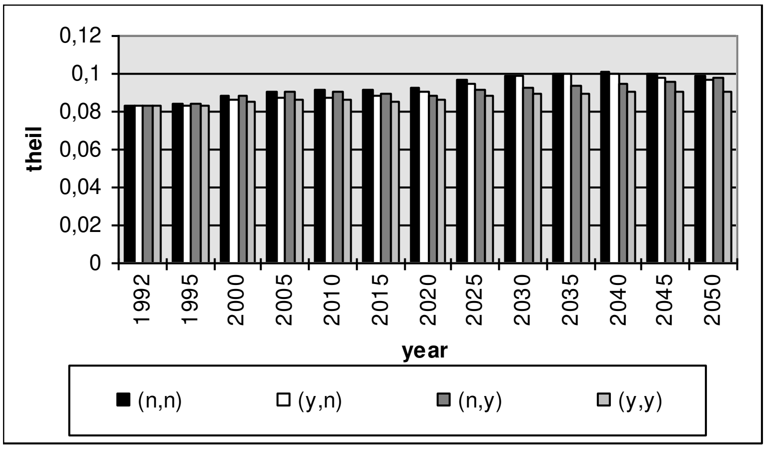
The effect of the implementation of the wage-barrier has is opposite to the effect to the effect of the partial linkage: in this case, the poverty among the non-retirees decreases whereas poverty among the retirees increases. This last effect is puzzling and goes against our intuition, for we would expect that a policy measure which has a limiting effect on the highest incomes of the retirees to have a poverty-decreasing effect. This would be true if the poverty line would have been one-half of the average income of the *retirees*. But it is not: the poverty line is based on the *sample-wide average income*. The increasing poverty among retirees turns out to be an artificial effect of changing average incomes. On the one hand, the average income of the retirees decreases. However, on the other hand, the average income of the sample as a whole decreases as well, but the decrease of the latter is less strong than that of the former. The result is that the average income of the retirees decreases relative to the sample-wide average, thereby causing poverty rates of the retirees to increase. Finally, note that the changes of these age-specific poverty rates indeed roughly cancel each other out.

Next, we turn to the distribution of income of the retirees. The Theil-coefficient T of the (equivalent) pension income is written as follows:

$T = \frac{1}{n}\sum_{i=1}^{n}\frac{y_i}{\bar{y}}\log\left(\frac{y_i}{\bar{y}}\right)$, where n is the sample size and y_i is the income of individual i (Cowell, 1995, p. 49) The value of the Theil-coefficient ranges between zero (in the case of complete equality of incomes) and log(n). Firstly, the future development of the sample-wide inequality of total income for the various simulation variants will be presented and discussed. Next, the within-group inequality for retirees and non-retirees will be presented separately, for the various simulation variants.

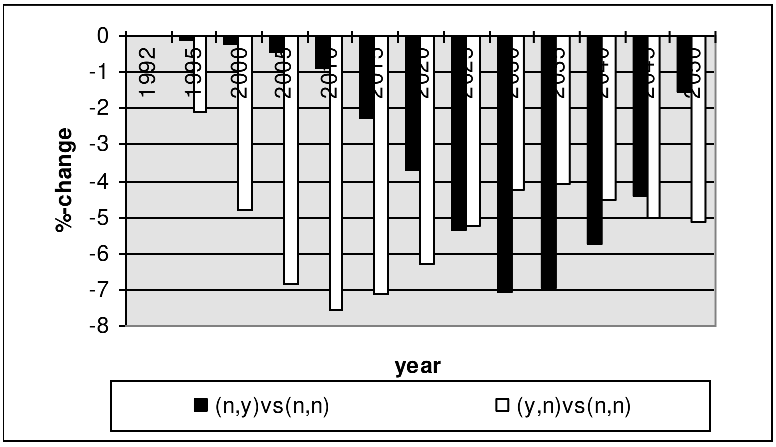
The effect of the various simulation variants on the sample-wide income inequality is given in the following figure 33.

Figure 33: sample-wide income inequality.



The Theilcoefficient is 0,0827 in the base-year 1992 and increases gradually up to about 0,1 in 2035 (base-variant). From then on, inequality remains stable. In the context of static microsimulation, this increase up to 2035 is closely related to the fact that older retirees (whose pension incomes are more widely distributed) get a higher weight in the calculation of the Theilcoefficient. It is however hard to establish a causal link: is it the decrease of income inequality over time in the past (for instance as a result of the social security system being extended), which causes the income inequality of younger age-groups to be lower than that of older age-groups and consequently of an increase of the income inequality as the older age-groups get a higher weight in the simulation process? Or is the opposite the case, and is it actually the case that pension income for all age-groups (including the retirees) has become more unequally distributed in the past, and in the future as well. The answer is hard to find, since it requires long-term information on past income inequality, which does not exists to my knowledge. The only information we found supports the second line of reasoning, and is that the income inequality for both the active population and retirees has been increasing between 1985 and 1992 (Cantillon, et. al., 1993, table 8 p. 13). Another possible explanation for the increasing inequality between 1992 and 2010 is not that the weight of the retirees in the calculation of the sample-wide inequality increases, but that the weights remain roughly the same whereas the income inequality within the group of retirees increases as a result of the uprating process. Which one of these above explanations is 'more true' than the other, will only become clear when looking at the within-group inequalities of retirees and non-retirees. This will be done below. For now, let us leave this, and consider the effect of either indexing the pension benefit and/or implementing the wage-barrier on the dispersion of pension income. The change of the income inequality as a result of the two partial simulation variants is given in figure 34.

Figure 34: effect of the partial linkage and the wage-barrier on income inequality.

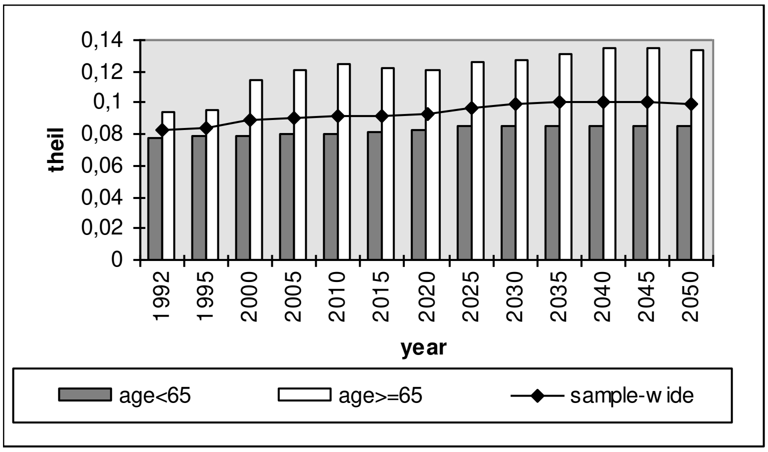


If we ignore the income-side of the model for a moment, we can a priori expect the linkage of the pension benefit to the course of wages to result in a lower income inequality, since the process of diverging average incomes is partly neutralised by this mechanism. This is clearly shown in figure 34: as a result of the partial linkage, income inequality decreases with at most 7,6%-point in 2010. However, as a result of the macroeconomic PAYG-identity, we know from figure 26 that the net incomes (which we are dealing with here) of the non-retirees grow at a lower speed than 2.25%. We saw in figure 26 that, in the long run, the average net income-level of the oldest category of non-retirees (which contribute to the system) becomes lower than that of two of three age-categories of retirees. Consequently, the inequality-decreasing effect of the partial linkage between the incomes of the retirees and the non-retirees will gradually become smaller in the long run. Again, this is shown in figure 34.

What is the a priori effect of the implementation of the wage-barrier to the dispersion of incomes? Earlier, on page 31, we concluded that the effect of the implementation of the wage-barrier to the dispersion of pension incomes was a priori ambiguous. However, we reasoned that the most likely effect would be a slow decrease of inequality at first, which would turn into an increase of inequality in the very long run. This is confirmed by the above figure: the ambiguity of this effect is reflected by its short-term weakness, as compared to the effect of the partial linkage. However, in the longer run, the positive effect on income inequality is very strong. At the end of the simulation period, this strong equalising effect is gradually -but not fully- neutralised, which also confirms our expectations.

Next, let us compare the development of income inequality of the retirees and the non-retirees. An important advantage of the Theilcoefficient is that it can easily be decomposed in within-group inequality and between-group inequality (i.e. within - and between the retirees and non-retirees). Let us start by considering the within-group inequalities. Figure 35 shows the within group-inequality of the retirees and the non-retirees together with the sample-wide income inequality in the case of no policy-measures.

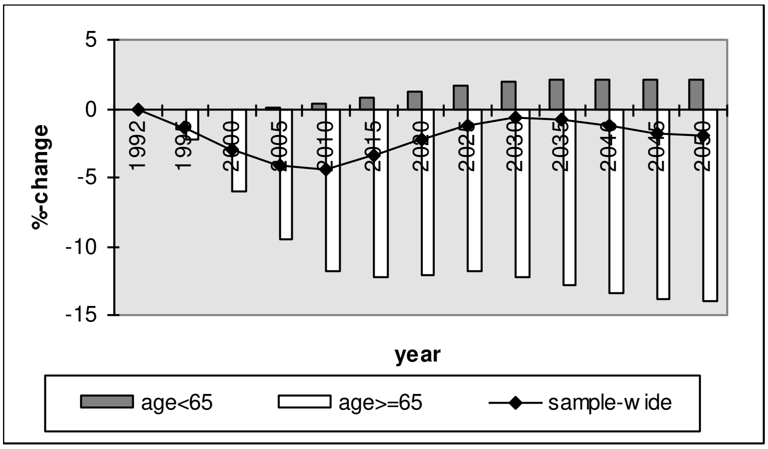
Figure 35: within-group- and sample-wide inequality: (n,n)-variant.



The development of sample-wide income-inequality can for a large extent (but not fully) be explained by looking at the within-group income inequalities. When considering figure 35, the first interesting thing is that the within-group income inequality of the group of retirees is higher than that of the non-retirees. This was already mentioned when discussing the sample-wide income inequalities in figure 33. Two possible reasons for the increase of the sample-wide income inequality where given when discussing figure 33. Both the increase of the weight of the retirees as compared to the weight of the non-retirees, as well as the increase of the within-group inequality of the retirees were possible causes for the increase of the sample-wide income inequality. Figure 35 suggests -but only suggests, since any more definitive conclusions can not be drawn- that it is merely the increase of the within-group inequality of the retirees which causes the sample-wide inequality to increase.

The next logical step is to consider how the within-group inequalities change as a result of implementing the partial linkage, the wage-barrier, or both. Again, the figures are expressed in percentage-changes. Figure 36 shows the effect of implementing the partial linkage.

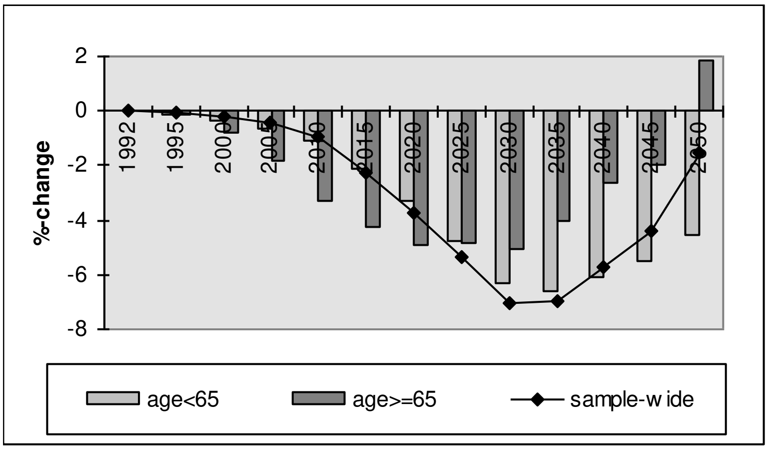
Figure 36: the effect of partially linking the pension benefit to the course of wages.



First of all, note that the introduction of the partial linkage between the income of the retirees and non-retirees does not have any effect on the income inequality of the non-retirees, since they are assumed to contribute the same extra amount of money to cover the cost of the partial-linkage. However, there are some retirees which are younger than 65 and whose pension income increases relative to the base-variant while they do not pay the extra contribution. So, for the category individuals younger than the retirement age, some individuals (the early-retired) get more while others (the non-retired) get less. As a result, income inequality goes up. By contrast and in line with earlier explanations, the inequality-decreasing effect of the partial linkage on the incomes of the retirees is very important and even going up to 13,9% in 2050.

The effect of the introduction of a partially-linked wage-barrier is shown in figure 37.

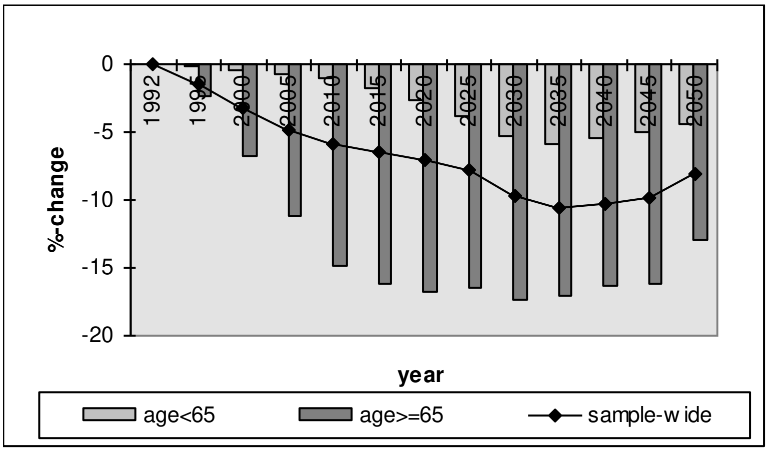
Figure 37: effect of the wage-barrier on inequality of income.



As a result of the introduction of the partially-linked wage-barrier, both the within-group inequalities as the between-group inequality decreases. What is interesting is that the inequality of the elderly (i.e. the retirees older than 65) seem to decrease less strong than the inequality of the retirees younger than the retirement age. A possible explanation could be that this last group of retirees generally has a higher (pension) income than the group of retirees older than the retirement age.

Lastly, figure 38 shows the combined effect of partially linking the wage-barrier as well as the pension-benefit to the course of wages on the sample-wide income inequality.

figure 38: joint implementation of the partially-linked wage-barrier and the partial linkage of pension benefit to the course of wages.



For both age-categories within the whole sample, the within-group inequality of income decreases, a development which of course can be explained by looking back at figures 36 and 37. This is especially the case for the main category of pension-beneficiaries, namely those older than the retirement age. For this category both partial effects point in the same direction, so the result of the joint implementation of the wage-barrier and the partial linkage of the pension benefit to the course of wages is unambiguously a decrease of the within-group inequality. This is also the case for the within-group inequality of the retirees younger than the retirement age. But we are considering all individuals, and the inequality of all individuals younger than the retirement age increases due to the increase of the contribution rate. This is the first effect. The second effect is the implementation of the wage-barrier and this results again in a decrease of inequality. It turns out that this second effect is stronger than the first effect. One can expect that those individuals who retire early receive a pension based on a relatively high wage, so that the implementation of the wage-barrier will have an important impact. Moreover, the years that individuals actually are early retired, are limited. Consequently, the effect of the imposition of the partial linkage of the pension benefit to the course of wages remains limited for pension beneficiaries younger than the retirement age.

So far, some rather simple and partial measures of wealth have been presented and discussed. But the simulation results of STATION allow for more complex analysis as well. As an example, let us consider a somewhat more advanced notion of wealth. Based on work by Dagum, Kakwani (Kakwani, 1986, p. 212-213) developed additive welfare functions for subgroups, using a subjective notion of 'envy'. Indeed, the wealth of a household is to a certain extent determined by the incomes of other households. In fact, this measure was explicitly developed to measure wealth disparities between to groups and bases on the idea that the wealth of households in one group is determined by comparison with the wealth of households in the other group[29]. Depart from the following notion of envy, represented by the variable *k*.

$$g(x,y) = x \text{ if } x \geq y$$
$$= x - k(y-x) \text{ if } y > x$$

So, "if the individual finds that the compared incomes [*y*,GD] are lower than his, then his welfare is given by his own income x. If, on the other hand, the compared incomes selected are higher than his, then the individual feels envious and loses welfare [...] proportional to the differences in incomes" (Kakwani, 1986, p. 198). The exogenous variable *k* denotes the impact of this income comparison on wealth: it can therefore be seen as the "parameter of envy". In what follows, let us apply the above, but then for those younger than 65 versus the individuals of 65 and older. Denote *g* the index number of the group, so that *g*=1 for the group of individuals younger than 65. Furthermore, denote a_g the proportion of individuals, μ_g the mean income, and *Gg* the Gini index of income distribution in group *g*. Moreover, define *j* as representing the 'other' group. So, if *g*=1 then *j*=2 and vice versa. Then the wealth W_g can be written as

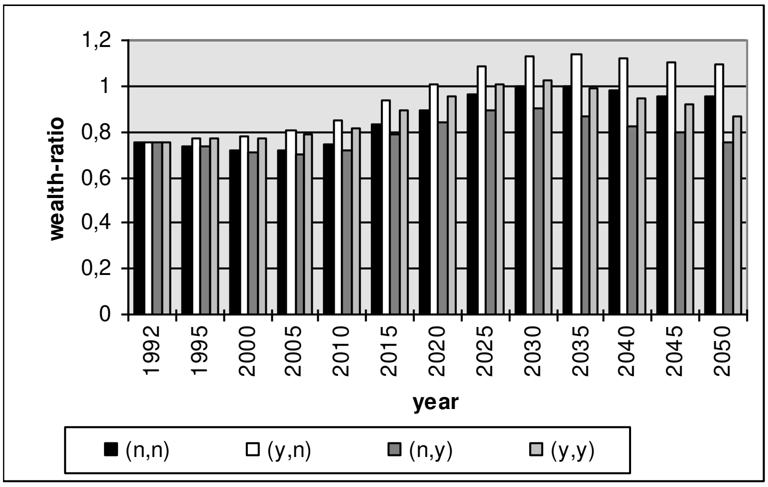
$$W_g = \mu_g - k\left[a_g \mu_g G_g + \frac{A}{2a_g} + \frac{1}{2} a_j \left(\mu_j - \mu_g\right)\right]$$

Equation 6

[29] Kakwani used this measure to consider the wealth difference between men and women. So, in this context, the wealth of a man was to a certain extent caused by the income of women.

where $A = \mu G - a_1^2 \mu_1 G_1 - a_2^2 \mu_2 G_2$ and μ and G are of course sample-wide values. Moreover, total wealth W can be written as the weighted sum of the wealth of the two groups: $W=a_1W_1+ a_2W_2$. However, this last variable W is not very interesting. What is of interest is the welfare-disparity as measured by the ratio of welfare of the two groups. The following figure 39 shows the development of the ratio of the Kakwani-index of wealth for the households of which the head is 65 years of older and that of the households of which the head is younger than 65.

Figure 39: Kakwani-wealth ratio with the envy-variable k=0.5.

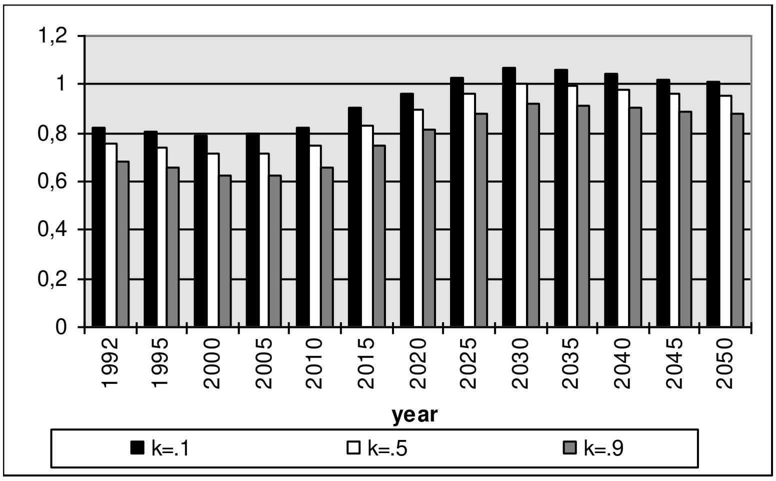


The initial value of the wealth ratio is .75 in the starting year 1992. Put differently, the wealth of the elderly is lower than the wealth of the non-elderly. Why is this? Simply because the elderly have more reason to be envious in the sense that the average income of the elderly is lower than that of the non-elderly. For the first period, the Kakwani-wealth ratio decreases to about .715 in the case of the base-variant. From then on, an increase sets in, resulting in the ratio reaching its maximum of almost 1 in 2030. After that, the ratio again decreases to .95 in 2050. As could be expected from earlier simulation results, implementing the partial linkage between the pension benefit and the course of wages results in a strong increase of the wealth of the elderly relative to the wealth of the non-elderly. In fact, relative to the Kakwani-wealth ratio in the base-variant, the wealth-ratio in this variant increases up to a factor 1.14 in 2010, which is when ageing really starts to kick in. After that, this factor remains more or less stable. The Kakwani-wealth ratio ends up a factor 1.15 higher than in the base-variant in the simulation-year 2050. So, as a result of the partial linkage of the pension benefit to the course of wages, the wealth of the elderly increases with 15% relative to the wealth of the non-elderly! This is caused by the fact that this linkage does not only increase the incomes of the elderly, but -via higher contributions- also decrease the income of the non-elderly. So, seen from the perspective of the elderly, not only do their incomes increase, but the 'reference incomes' (which the envy is based upon) decrease as well. As a result, the wealth of the elderly increases. Of course, this is accompanied by a decrease of the wealth of the non-elderly, which explains the strong effect on the ratio of the two. It is equally clear that the implementation of the wage-barrier will result in a decrease of the Kakwani-wealth ratio relative to the base-variant. The reason is that this wage-barrier will limit pension-benefits and decrease contributions made by the non-elderly, again all relative to the base-variant. So, the wealth of the elderly will decrease whereas the wealth of the non-elderly will increase, resulting in a decrease of the above-described ratio. This decrease is rather modest for the largest part of the simulation period. In 2025, the Kakwani-wealth ratio will have decreased to .93 of its value in the base-variant. From then on, the decrease becomes stronger, resulting in a value being about .79 of its original value in 2050.

Given the above results, it may again come as no surprise that the combined effect of both measures on the Kakwani-wealth ratio is limited. Moreover, it is not consistent in the sense that the direction of the effect changes over time. At first, the joint implementation of both measures results in an increase of the Kakwani-wealth ratio relative to its value in the base-variant. This difference reaches a maximum of about 9% in 2010. After that, the ratio again decreases and reaches a minimum of .91 in 2050. It appears that the positive effect of the partial linkage of the pension benefit on the wealth ratio comes before the negative effect of the implementation of the wage-barrier. This is in line with what we saw when discussing the effect of both measures on the dispersion of income (figure 34), and can be explained by the somewhat ambiguous short- and middle-term effect of the wage-barrier on the dispersion of income.

In the above simulation results, the exogenous parameter of envy *k* was set equal to .5. What would be the effect of increasing or decreasing the enviousness of the individuals? In the following figure 40, the simulations for the base-variant where again calculated, but then for *k* being equal to .1 (low envy) and .9 (high envy).

Figure 40: the effect of the degree of envy on the Kakwani-wealth ratio.



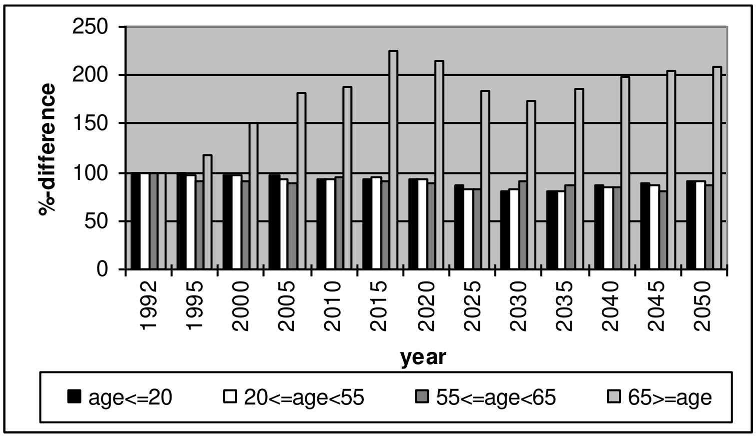
It is clear that the Kakwani-wealth ratio decreases as the envy-variable *k* increases. As said, the wealth ratio is below 1 because the wealth of the elderly -whose income is lower than that of the non-elderly- forms the numerator of this ratio. As the elderly become more envious to the income of the non-elderly, their wealth will by definition decrease. The wealth of the non-elderly will however not be affected, or at least not as strong as that of the elderly, which follows directly from the equation introducing *k*.

3.4. Changing the point of view: what would happen if certain measures were *not* taken?

In the first three sections of this third chapter, it was considered what the effects of two policy measures on incomes, poverty, income-inequality and welfare would be. These measures were the partial linkage of ongoing pensions to the course of wages and the implementation of a partially-linked wage ceiling. The point of departure was the situation in which there would be no linkage and no lagging wage-ceiling. It could be interesting to take the fourth simulation variant (y,y) as the point of departure, for it resembles the closest to the 'working hypothesis' of the Federal Planning Agency. This working hypothesis is based on the historical situation in the period 1969-1991. However, from about 1981, the wage-ceiling was no longer adjusted to the course of wages. So, let us consider the fourth variant, and see what happens to the position of the elderly if one or both of the well-known measures is *not* taken. Before doing so, let us briefly review what happens if both measures are taken. The partial linkage of pensions to the course of wages results in an increase of the contribution-rates for the non-retired. So, there is increasing redistribution of income from the non-retired to the retired. As a result of maintaining the partially-linked wage-ceiling, the highest pension-benefits no longer grow at the same rate as the other pension benefits. So, the contribution-rates for the non-retired relatively decrease, and there is a certain redistribution from the retired to the non-retired.

What would be the result of abolishing the partial-linkage of ongoing pensions with the course of wages on poverty? This can be seen from figure 41.

Figure 41: effect on abolishing the partial linkage between wages and pensions on poverty.



It is clear that not linking pensions with wages causes important redistribution of income from the retired to the non-retired. The contributions of the non-retired decreases, with the result that poverty among individuals in this category decreases. The poverty among retired of course increases strongly. At most, it increases with 116% in 2015. It can be seen that the abolishment of the wage-ceiling on poverty are less extreme.

Figure 42: effect on abolishing the wage-ceiling on poverty.

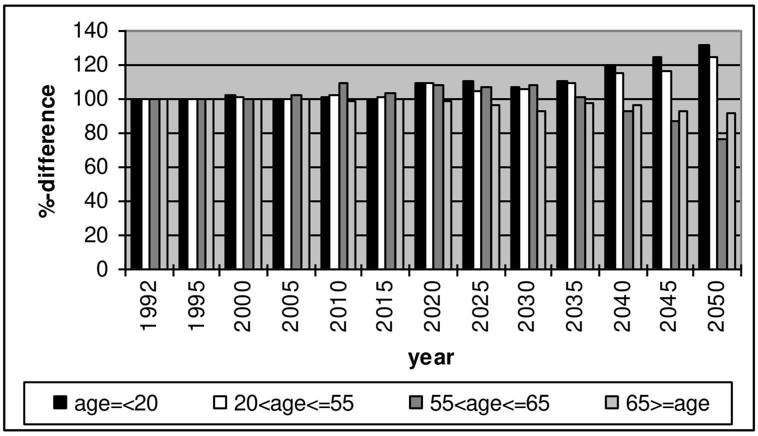


Figure 42 shows that the abolishment of the wage-ceiling does not have an effect in the short run. However, in the long run, it causes poverty among the non-retired to increase whereas the poverty among the retired decreases. This is the consequence of the fact that this measure causes contributions to decrease relative to their development otherwise. So, the abolishment of this measure increases contributions, thereby causing income redistribution from the non-retired to the retired.

To see what would happen to income inequality if these policy measures would not be introduced, it is sufficient to change the directions in figures 35, 36 and 37. If neither of both measures are implemented, income inequality among both retired and non-retired would increase. For the retired, the increase will be drastic: to at most 21% in 2030. The inequality among non-retired will increase with at most 6% in 2035. Sample-wide income inequality would increase with at most 11% in 2035 as well.

The effects of abolishing these measures on the Kakwani-wealth ratio are shown in figure 43. It is clear that abolishing the partial linkage between wages and pensions will result in a decrease of the wealth of elderly relative to non-elderly. This decrease will come quite fast, and the wealth of the elderly will in 2010 be 12% below its level if the measure would have been implemented. Likewise will not introducing the wage-ceiling result in an increase of the wealth of the retired relative to the non-retired. This effect will only emerge in the long run, but will be considerable. In 2050, the wealth of the non-retired would be no less than 26% higher as compared to the situation if the measure was actually taken.

Figure 43: effect on abolishing the partial linkage between wages and pensions and/or the wage-ceiling on the Kakwani-wealth ratio.

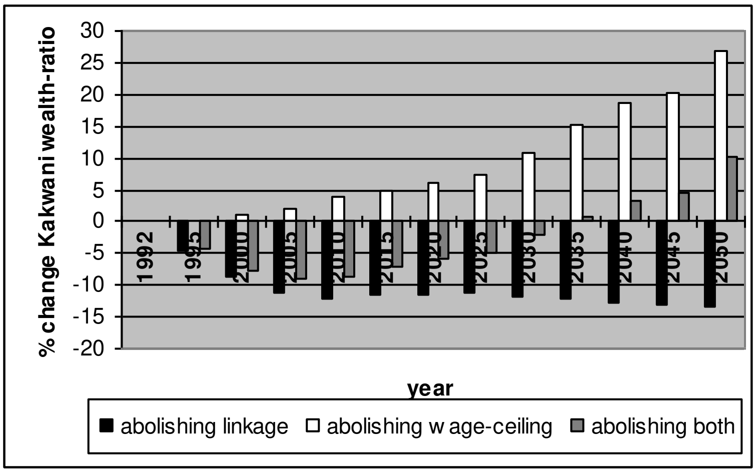


Figure 43 shows as well that, if both measures would not have been taken, then the wealth of the retired would at first decrease relative to that of the non-retired, with a maximum of a bit more than 9% in 2005. After that, it would increase again, to at most 10% in 2050.

Chapter 4: criticisms, discussion and conclusions.

4.1. Criticisms and discussion.

In the first chapter, the problem definition, as well as need for developing a microsimulation model was described. Then, in the second chapter, the static microsimulation model STATION was presented and discussed. The third chapter then contained an extensive discussion of the simulation results. Before turning to the summary and conclusions of this study, however, it seems useful to reconsider one last time the model which has been developed, and to see in which ways it could be improved and -if so and if possible- how this could be done. Along the way, some arbitrary decisions had to be made, and it would not seem right not to highlight these as well. Therefore, in this one-but-last section, some criticisms will be brought forward and discussed. One could see at least some of these criticisms as a plea to try to solve these problems in the future.

First of all, why not developing a dynamic microsimulation model instead of a static one? After all, the former are more advanced than the latter nowadays, and Belgium is one of the few countries in Europe which does not have such a model.

That all being true, there still was no reason to develop a dynamic model *given the current research problem*. If it comes to the relative income position of the elderly, or inequality and poverty based on annual incomes, most other researchers have chosen static models as well (for a more elaborate discussion, see Dekkers, in Becker et.al, 2000). The reason for this is that the development period for a dynamic model is much longer than for a static model. In fact, for just one research problem, the effort of developing an entire dynamic model is just not worth wile, both in terms of development time and financial sacrifice. Moreover, the development of a dynamic model requires the availability of specific data on transition probabilities. It is questionable whether much of this data exists or, which is the same in the end, whether we could have access to it.

A more practical problem involved in the above model concerns the use of income. As said in the second chapter, the household income of the individuals forms the point of departure. However, both the transformed weighting variable as well as the uprating process is based on individual information, such as age, labour market status (retired or not retired) and household-category. Consequently, the resulting simulation results are no longer based on a household-income concept which is equal for all members of the household. This is not very important as far as the reweighting scheme is concerned (as extensively explained in section 2.4.4.), and there were some good reasons for it too. However, the choice for an individual upgrading process, and thereby allowing the incomes of individuals in the same household to differ over time, remains somewhat arbitrary. Either this, or the uprating process for all individuals in a household should have been based on the information of only one individual in that household, notably the head of the household. This would have meant that the age distribution of other members of the household would have been completely ignored. So, the household income of an 50-year old head with a 40-year old wife would have been uprated exactly the same way as the household income of a 50-year old head with a 50-year old wife. This would ignore an important amount of information. So, a subjective choice had to be made, and it was chosen to let the uprating process use individual information, thereby ‘sacrificing’ the equality of household income for all members of the household. The most likely effect of this decision will have been an increase of income inequality, however without disturbing the effects of the simulation variants.

Another possible criticism again concerns the uprating scheme which the model uses is that the macroeconomic environment is very rough and simple. First of all, it is not possible to change the growth rate of wages nor the linkage between wages and pension benefits along the way, at least not without the model behaving as if the new information has been applied from 1992 on. Secondly, past macroeconomic changes are not taken into account, at least those which are not already implicitly incorporated in the 1992-income distribution. To clarify this, let us consider an example. During the sixties, wages rose considerably and these wage-increases have a hidden but probably important effect on the current level of pensions. Now if we use current uprated pension benefits as a proxy to future pension benefits, we implicitly assume that future retirees will have the same wage-history: their own 'sixties', so to say. Moreover, we assume that the past pension scheme is the same as in the one in the future. Of course, this is a problem. However, could one argue, it is a problem which all empirical models face. All information underlying any economic model is by definition ex post information. For instance, in dynamic microsimulation models, future wages are determined by a wage-function, which is estimated on historical information. That is true, but in a dynamic microsimulation model, future pensions are based on the wage-function, which is estimated using wage-information of 1992. However, in a static model, future pensions are based on current (1992-) pensions, which are based on much earlier wages, as well as the past development of the pension system. So, the 'time-lapse' is larger in the case of a static model than in the case of a dynamic model. This increases the risk that time-specific circumstances pollute the simulation results.

4.2. Summary and Conclusions.

It is beyond any doubt that the fact that the Belgian population is ageing causes concern for the sustainability of the pension scheme. Numerous models have been developed to analyse the relation between ageing and the pension system. The Belgian Federal Planning Bureau has developed a model, PENSION, which shows this relation. This model includes some assumptions on economic growth, the linkage between contributions and benefits and the existence of a pension-ceiling. Whether or not the conclusions drawn are valid or not, is not the point of interest here, but the model fails to show the distributional effects of these assumptions. However, when taking a decision on which potential measure to adopt, politicians should know about the effect of these measures on poverty and income inequality.

The static microsimulation model STATION is therefore complementary to the model of the Federal Planning Bureau. It is designed to show distributional effects of changing demographic circumstances, as well as the uprating of monetary variables. The model disentangles four simulation variants. The first one is that there is no linkage between the pension benefit and the course of wages and no wage-barrier in the determination of the future pension benefit. The second and third simulation variants partially adopts the linkage and the wage-barrier, respectively. Lastly, the fourth variant adopts both the partial linkage and the wage-barrier. Of course, the model is completed with a so-called PAYG-equation, assuring the equality between extra benefits and contributions.

The first simulation results concern the comparison of average household incomes of various categories of households (where the age of the head of the household typically is the relevant variable). Figure 24 on page 59 shows that "neither the partial or combined implementation of the (partial) linkage of the pension benefit and the (partially-linked) wage-barrier is a Pareto-improvement". Maybe, but this is given a rough utilitarian welfare-concept and moreover ignores the distributional effect of these measures. Figures 26 and 27 show that especially the partial linkage of the pension benefit to the course of wages has got a strong effect on average incomes; especially that of the oldest retirees increases very considerably. This pattern shows for the development of poverty rates as well, as could be expected. Figure 29 shows that poverty among the retirees will increase strongly in the first half of the simulation period. After 2010, it will decrease again and will converge to the (gradually increasing) poverty rate among households of which the head is between 55 and 65. The implementation of the wage-barrier clearly has a very limited effect on poverty, but figure 31 shows that the effect of the partial linkage of pensions to the course of wages results in a strong decrease of the poverty among the retirees (eventually down to 50 percent!) causing the general poverty rate (figure 30), to decrease as well. It may come as no surprise that the sample-wide income inequality, represented by a Theil-index, increases over time, and that both simulation variants have a inequality-decreasing effect (figure 32). However, first of all, the effect of these measures is less strong than in the case of the poverty rates (though considerable), but more importantly and as opposed to what was the case with the poverty rates, the effect of the implementation of the wage-barrier now is at least as strong as the effect of partially linking the pension benefit to the course of wages. However, it is less straightforward, in the sense that it is difficult to say a-priori what the effect of this measure on inequality will be. Figure 32 shows that income inequality decreases at first, but that this effect does not last. When considering the within-group inequalities, the only general conclusion to be drawn is that as a result of both measures, the income inequality among households with a head of 65 and older, will decrease considerably (see figures 36 and 37). For the other households, the situation is ambiguous, but a combined implementation of both measures will result in a lower income inequality as well, as figure 38 shows. Lastly, the Kakwani-wealth ratio, which defines wealth not only to be a function of ones income, but also a function of higher incomes in the 'other' category. To this goal, an 'envy-variable' is introduced. Figure 39 shows that especially the introduction of the partial linkage of pension benefits to the course of wages results in a considerable increase of the wealth ratio, meaning that the wealth of the elderly shifts closer to the wealth of the non-elderly. Moreover, it shows that the introduction of the wage-barrier will result in a loss of welfare of the elderly relative to the young. This of course is not unexpected, since this measure will limit the costs of the pension benefits paid out to the retirees. Lastly, figure 40 shows that an increase of the 'envy-variable' results in a lowering of the welfare of the elderly relative to that of the non-elderly.

Now what conclusions can there be drawn from this study? First of all, the simulation results show that, both poverty and income inequality will increase. An important determinant of this is the uprating of income and pensions. Secondly. the impact of the partial linkage of the pension benefit to the course of wages and the implementation of the wage-barrier on poverty, income inequality and welfare, is quite important. When considering the effect of these measures on poverty among the elderly, especially the implementation of the partial linkage between wages and pensions has a strong and unambiguous poverty-reducing effect. This reduction in poverty is accompanied by a unambiguous decrease of income inequality, which does not come as a surprise. The effect of implementing the wage-barrier on poverty rates is negative, but much smaller. Moreover, the effect on income inequality was a priori unclear, but turns out negative. So, as a result of the implementation of the wage-barrier, income inequality decreases as well, though its effect is limited, both in time and in magnitude. But that is not all, of course, as a result of the implementation of both policy measures, there is an implicit income redistribution: the linkage of the pensions to the course of wages causes a redistribution of income from all non-retirees to all retirees (via the PAYG-equation). However, the implementation of the wage-barrier reduces the highest pension benefits which results in a relative decrease of the contributions of the non-retired. So, the effect of this measure is an income redistribution from rich pensioners to all non-retirees. Combining these measures therefore implies a redistribution of income from the non-retired and the rich retirees to the non-rich retirees. Even though this might not be Pareto-optimal, it surely seems a good thing to do, from the distributional point of view. This is confirmed by the Kakwani-wealth ratio depicted in figure 39: both measures to a certain extent neutralise each other, so that the costs (and effects of these costs on poverty) are borne by those who can carry them.

Literature

Aaron, H., 1977, *Demographic Effects on the Equity of Social Security Benefits*, in: Feldstein, M., Inman, R.(eds), *the Economics of Public Services*, MacMillan, New York, pp. 151-173.

d'Alcantara, G., Wardenier, R., 1997*, De Socio-professionele Statuten uit de Sociale Zekerheid in België*, Ministerie van Sociale Zaken, Volksgezondheid en Leefmilieu, Garant, Leuven.

Barry, N., 1990, *Welfare*, Concepts in Social Thought 1990:1, University of Minnesota Press, Minneapolis.

Becker, H., 1994, *Generaties en hun Kansen*, fourth edition, Meulenhoff, Amsterdam.

Bengtsson, T., Kruse, A., 1993, *Impacts of Demographic Changes on an Economic Growth Sensitive Pension System: the Case of Sweden*, Paper presented at the 7th annual Conference of the European Socity for Population Economics (ESPE), Budapest June 3-5, 1993.

Blackburn, M., 1998, The Sensitivity of International Poverty Comparisons, Review of Income and Wealth, vol 44, no 4, pp. 449-472.

Bolhuis, E., Vossers, W., 1986, *Pensioenfondsen in Nederland*, no 38, Centraal Planbureau, 1986.

Bos, D. (ed), 1993, *Economics in a Changing World*: Proceedings of the 10th annual world congress of the International Economic Association, Moskow, 1993, Vol. 3.

Boulanger, P., A. Lambert, P. Deboosere and R. Lesthaeghe, 1994, *la Formation des Familles: étude Prospective*, programme de recherche en sciences sociales, service de programmation de la politique scientifique.

Cantillon, B., Houvenaghel, I., Meulemans, B., 1994, *De Pensioenkwestie vanuit Verdelingsstandpunt*, Belgisch Tijdschrift voor Sociale Zekerheid, 1994, vol. 36, no 3, 1994, pp. 777-808.

de Callatay, E. , Turtelboom, B. , 1996, *Pension Reform in Belgium*, IMF working paper WP/96/74 July, p.32.

Cantillon, B., Marx, I., Proost, D., Van Dam, R., 1993, *Sociale Indicatoren*, CSB-berichten D/1993/6104/10, Centrum voor Sociaal Beleid, Universiteit Antwerpen - UFSIA.

de Cock, J., 1984, *De Pensioenproblematiek: Elementen bij een actueel debat*, Centrum voor Politieke, Economische en Sociale Studies (CEPESS), Brussel.

Creedy, John (ed.), 1995, *The Economics of Ageing*, Elgar Reference Collections: International Library of Critical Writings in Economics, Vol. 51, Aldershot, U.K.

Cowell, F., 1995, *Measuring Inequality*, 2nd Ed., LSE handbooks in Economics, Prentice Hall, London.

Dekkers, G., Nelissen, J., Verbon, H., 1995, *Intergenerational Equity and Pension Reform: The Case of the Netherlands*, Public Finance, vol. 50, pp. 224-245.

Dekkers, G., 1998, *Het Nederlandse Pensioenstelsel vanuit Belgisch Perspectief*, Belgisch Tijdschrift voor Sociale Zekerheid, vol 40 no 1, pp 35-59.

Dekkers, G., 2000, *Microsimulation*, in Becker, H., F. Vanclay, C. Wolf (eds), 2000, *Conceptual and Methodological Advances in Social Impact Assessment*, forthcoming. Chapter 5.

van Eeckhoutte, W, 1998, *Sociaal Zakboekje* 98-1, Kluwer/Samsom.

Festjens, M.-J., 1997, *De Pensioenhervorming*, Federaal Planbureau Planning Paper 82, November.

Harding, A., 1993, *Lifetime Income Distribution and Redistribution. Applications of a Microsimulation Model*, Contributions to Economic Analysis 221, North-Holland, Amsterdam.

Harding, A.(ed), 1996, *Microsimulation and Public Policy*: selected papers from the IARIW Special Conference, Canberra, 5-9 December 1993, Elsevier, Amsterdam.

Huijser, A., van Loo, P., 1986, *Vergrijzing, Pensioenen en Contractuele Besparingen*, Monetaire Monografieën nr 5, De Nederlandsche Bank n.v., Kluwer Deventer.

Heinrich, G., 1998, *Changing Times, Testing Times: a Bootstrap Analysis of Poverty and Inequality Using the PACO database*, CERT-working paper, Heriot-Watt University, March 1998.

Jackson, W., 1992, *Population Ageing and Intergenerational Conflict: a post-Keynesian view*, Journal of Economic Studies, Vol. 19, no 2, pp. 26-37.

Jackson, W., 1998, *The Political Economy of Population Ageing*, Edward Elgar Publishing Ltd., Cheltenham (UK), Northampton MA (US).

Jansweijer, E., 1996, *Gouden Bergen, Diepe Dalen: de inkomensgevolgen van een betaalbare oudedagsvoorziening*, Wetenschappelijke Raad voor het Regeringsbeleid, Sdu Uitgevers, Den Haag

Kakwani, N., 1980, *Income Inequality and Poverty*, Oxford University Press, New York/Oxford, etc..

Kakwani, N., 1986, *Analyzing Redistribution Policies: a study using Australian data*. Cambridge University Press, Cambridge.

Kalton, G., 1983, *Compensating for Missing Survey Data*, Research report series, Institute for Social Research, University of Michigan.

Kingsland, L., 1982, *Projecting the Financial Condition of a Pension Plan using Simulation Analysis*, the Journal of Finance, vol 37, p. 577-584.

De Lathouwer L., (1996) *Microsimulation in comparative social policy analysis: a case-study of unemployment schemes for Belgium and the Netherlands*. - In: Harding (ed.), 1996, Amsterdam, Elsevier, p. 69-91.

Lesthaeghe R., Meeusen, W., Vandewalle, K, 1998*, Eerst Optellen, dan Delen; Demografie, Economie en Sociale Zekerheid*, Garant Leuven/Apeldoorn.

Nationaal Instituut voor de Statistiek (ministerie van Economische Zaken), 1992, *Algemene Volks- en Woningtelling op 1 maart 1991*, deel 1B: Bevolkingscijfers, NIS, Brussel.

Belgian National Institute for Statistics and the Federal Planning Bureau, 1993, *Bevolkingsvooruitzichten 1992-2050*, Brussel.

Merz, J., 1994, Microsimulation - A Survey of Methods and Applications for Analyzing Economic and Social Policy, Universität Lünebürg, Discussion Paper no 9, June 1994.

Nelissen, J., 1998, *Annual versus Lifetime Income Redistribution by Social Security*, Journal of Public Economics, vol. 68, nr 2, pp. 223-250.

Neyt, P., 1993, *Het Belgisch Pensioenstelsel: een driepijlerstelsel*? in: *Ouderen in Solidariteit*. Vlaams Welzijnscongres 1993, Kluwer Editorial pp. 353-372.

Osberg, L., Xu, K., 1997, *International Comparisons of Poverty Intensity: Index Decomposition and Bootstrap Inference*, Paper presented at the 51st session of the International Statistical Institute, Istanbul, Turkey, August 22, 1997.

Proost, D., R. van Dam, K. Van den Bosch, 1996, *Valideringsstudie van twee Belgische Inkomenssurveys*, CSB-berichten D/1996/6104/07, Centrum voor Sociaal Beleid, Universiteit Antwerpen - UFSIA.

Quinn, J., 1997, *Economic Implications of an Ageing Society*, Review of Income and Wealth, vol. 43, no 4, December 1997, pp. 511-520.

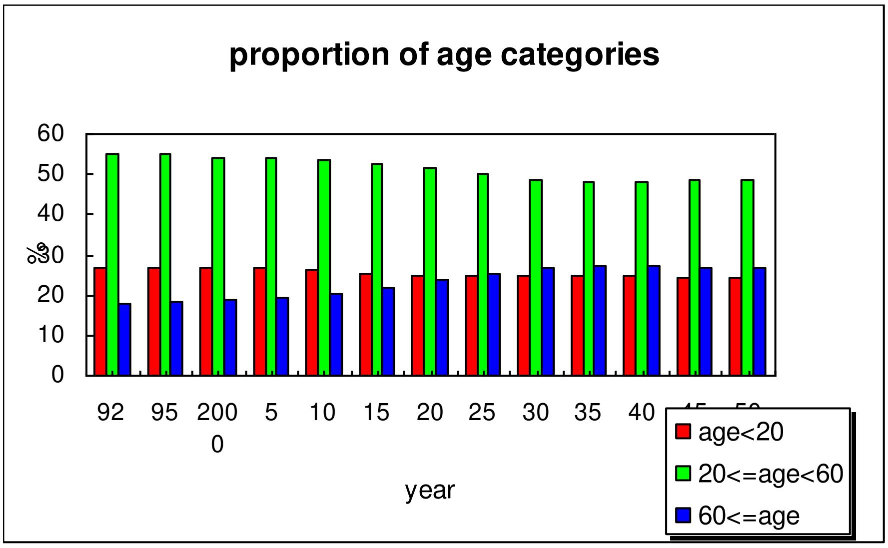
Shorrocks, A., 1995, *Revisiting the Sen Poverty Index*, Econometrica, vol. 63, no 5, september 1995, pp. 1225-1230.

Slesnick, D., 1998, *Empirical Approaches to the measurement of Welfare*, Journal of Economic Literature, vol. 36, nr. 4, december 1998, pp. 2108-2165.

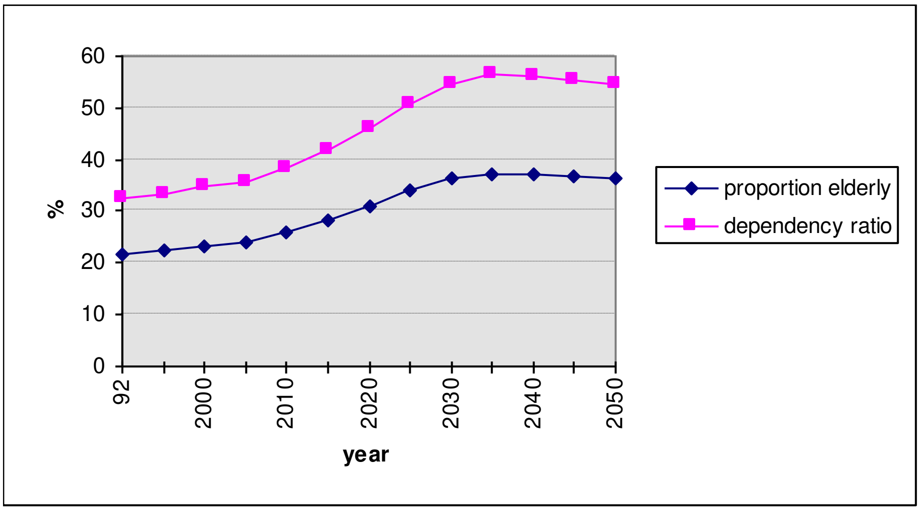
Verbist, G., Van den Bosch, K., Cantillon, B., 1999, *MISIM: microsimulatiemodel van personenbelasting en parafiscaliteit*, Centrum voort Sociaal Beleid, CSB berichten, D/1999/6104/04.

World Bank, 1994, *Averting the old-age crisis: policies to protect the old and promote growth*, World Bank Policy Research Report, 1994:1, World Bank, Washington D.C.

nog wat 'nagekomen' grafieken waarvan ik me afvraag of ze er in moeten:



bron: bevolk.wk4: M39



bron: bevolk.wk4: U9