WP 15.1
16 October 2013

UNITED NATIONS STATISTICAL COMMISSION and ECONOMIC COMMISSION FOR EUROPE

STATISTICAL OFFICE OF THE EUROPEAN UNION (EUROSTAT)

Joint Eurostat/UNECE Work Session on Demographic Projections
organised in cooperation with Istat
(29-31 October 2013, Rome, Italy)

Item 15 – Beyond population projections by age and sex: Inclusion of additional population characteristics

Projecting Inequality: The Role of Population Change

Ingrid Schockaert, Vrije Universiteit Brussel
Rembert De Blander, University of Leuven
André Decoster, University of Leuven
Patrick Deboosere, Vrije Universiteit Brussel

Projecting Inequality: the Role of Population Change

Ingrid Schockaert*, VUB, Dept. of Sociology (Ingrid.Schockaert@vub.ac.be)
Rembert De Blander, KU Leuven, Dept. of Economics (Rembert@DeBlander.eu)
André Decoster, KU Leuven, Dept. of Economics (Andre.Decoster@kuleuven.be)
Patrick Deboosere, VUB, Dept. of Sociology (patrick.deboosere@vub.ac.be)

30th September 2013

1 Introduction

In this paper, we analyse the evolution of inequality in Flanders over the next twenty years. We build on our previous work focussing on the impact of population change on policy indicators, as compared to role of economic transformations (De Blander et al., 2013). One of our findings was that between 2011 and 2031, the gini-coefficient will slightly decrease and that indeed population change is the driving force behind this process, counteracting the wage and benefits evolution that would otherwise have increased inequality. In the present paper we aim to disentangle the demographic elements most affecting inequality, and single out socio-demographic groups moving to the lower or upper end of the distribution in order to better understand, ex ante, the confluence between population change and inequality.

We focus on the ageing of the population and concomitant modifications in household structure and educational attainment. Based on multistate projections, over the next twenty years Schockaert and Surkyn (2012) foresee a pronounced growth of individuals over 65 and a net decrease in the adult population at working age. Parallel to the ageing process, from one projection year to the next, family formation of the current younger generations are transferred to successive age groups. Younger generations currently marry less and later, and they have less children than older generations did. divorce rates have become higher. Consequently, single headed households now mainly restricted to the young, become common among all ages. Parallel to ageing and changes in household position, De Blander et al. (2013) foresee a considerable advancement in educational attainment: from one projection period to the next, the current high educational levels of the young are successively transferred to older generations. Consequently, in 2031 the generational gap in education is firmly reduced and, with the exception of the very advanced age groups, women are outperforming men.

We thus join a strand of recent research aiming at quantifying the role of demographic change for inequality (Peichl et al., 2012; Burtless, 1999; Breen and Salazar, 2010; Daly and Valletta, 2006). In this recent literature a positive relationship is found between the prevalence of single headed households, especially single motherhood, and income inequality (Kollmeyer, 2012). The gender pay gap combined with the impossibility to pool income, in addition to lower educational levels of single mothers, are advanced as explanations for this relationship. Nevertheless, the negative impact of single motherhood is found to be mainly restricted to the United States as the strong European welfare state would mitigate the impact (Esping-Andersen, 2007). Educational advancement, leading in the first place to higher wages and labour force participation, is often found to increase inequality; combined with assortative mating, it leads to increased income disparity between highly educated working couples on the one hand, and lowly educated couples with low work intensity or single headed households on the other (Esping-Andersen, 2007). However, the larger the group of educated and the lower educational disparity, the smaller the role this factor will play in the determination of income inequality. In the long run, we well might expect the rise in educational attainment to temper inequality.

In contrast to former studies, we investigate population change and inequality prospectively. This is important from a policy point of view: for social policy aiming at guiding societal evolution to go beyond ad hoc answers to structural changes, it needs to be informed about future tendencies. Therefore, we build on the methodology established in De Blander et al. (2013), based on micro-simulation. A microsimulation model (*MSM*) is essentially a *forecasting* device that simulates aggregate and distributional effects of a policy - or

*Contact author.

another kind of change, such as population change - by applying them to a representative sample of individuals or families, subsequently adding up the results across individual units using population weights (Martini and Trivellato, 1997; Bourguignon and Spadaro, 2006). Population change is implemented by static reweighting of the EU-SILC 2008 data on the basis of the before mentioned population projections (Schockaert and Surkyn, 2012; De Blander et al., 2013). This reweighting procedure allows measuring household income distributions for 2011 and for 2031 congruent with the population composition of projections by age, sex, household position and education. The results in this paper were obtained using the microsimulation model *Euromod*.

To disentangle the impact of ageing and its covariates on inequality, we combine several methods. first, a decomposition of the Theil index and poverty rate allows to estimate the contribution of changes in each demographic *subgroup* by age, household position and education, to change in inequality. Secondly, we *isolate* the effect of each population *covariate* - age, gender, household position and education - constructing counter-factual income distributions. This involves reweighting the sample 2011 such that the distribution of education levels becomes identical to the sample at 2031. Comparison of the counter-factual with the actual distribution in 2011 reveals the effect of changing education, while the difference between counter-factual and the actual distribution at 2031 is a residual effect. Applied to the mean of a distribution, the second method is known as the Blinder-Oaxaca decomposition. In the context of distributional analysis it has been applied by DiNardo et al. (1996), Hyslop and Maré (2005), Handcock and Morris (1998), to name but a few. Handcock and Morris (1999) developed the relative distribution method. This method allows to move beyond the analysis of the means alone, to decompose the co-variate effects on the whole income distribution in detail. Moreover, the method detects subtle changes in the upper and lower tails of the distributions and indicates how this is related to changes in population attributes. Finally, based on quantile regression, we identify the changing role of age, gender, household position and education in determining individual household income. This analysis will give indications about the changing income position of individuals with different characteristics.

2 Methodology

Our methodology consists of two parts. First, we project the population data as found in the census 2001 to the year 2011 and 2031. We link these projections to the *EU-SILC* 2008 data to forecast household income to the same years. The way this is done is explained in Subsection 2.1. Secondly, we analyse changes in income distribution and inequality between 2011 and 2031 using indicators such as the Theil index and poverty rate on the one hand and relative distribution methods on the other. Decomposition methods will shed light on the impact of the aspects of population change. Quantile regression indicates the changing role of gender, age, education and household position in determining household income. Note that this analysis is restricted to individuals that have finished their formal educational training. The methods are explained in Subsection 2.2.

2.1 Preparing the data: Lipro-projections and static ageing

The method in this paper strongly builds on the proposal of De Blander et al. (2013), involving:

1. a Lipro population projection (Lifestyle Projection) (Van Imhoff and Keilman, 1991) at five-year intervals, for the next twenty years (from 2011 up to 2031), by age, sex and household position and by age, sex and educational attainment.

2. a calibration of the *EU-SILC* 2008 data to the distribution of the different population subgroups obtained by the demographic projections. This procedure is called static ageing.

Lipro-projections estimate population structures prospectively by multiplying the density of the initial observed population state vector (baseline vector) with a transition matrix to obtain the density of the state vector in the next period. This way we recursively project the population one period further. The transition rate matrix indicates the probability to transit from one household position (educational level) to another and also includes death and emigration rates *from* each household position (educational level) as well as births and immigration *to* each household position (educational level). The initial rate matrix is externally obtained from observed data. Rate matrices for consecutive projection periods are estimated following plausible scenarios about the evolution of fertility, mortality, migration, household formation processes and educational behaviour.

The population as found in the 2001 Census data served as the baseline state vector for both projections. For the projection by household position, the population is broken down by age and sex and 12 "Lipro household

positions": children of married and unmarried couples, children in lone parent households, married and unmarried couples with or without children, single households, lone parents, non-related family members, members of collective households and an "other category" [1]. For the projection by educational attainment level, the state vector comprises age, sex and nine educational levels: individuals can be still in school or have finished school with a diploma of primary education, lower secondary education (general, technical and professional), higher secondary education ((general, technical and professional) and higher education. Initial rate matrix was obtained from linked 2001 census and Register Data.

For each consecutive five year period up to $2026-2031$, the transition rates are adapted using scenarios based on recent prognoses about the evolution of fertility, mortality and migration(Studiedienst Vlaamse Regering, 2011): we assume that the recent revival of fertility in Flanders will continue up to the period $2016-2021$ to go down again afterwards; ; life expectancy will continue to increase, a little faster for men than for women; we assume international immigration to increase up to the $2016-2021$ period; emigration increases linearly with about 20% over the whole projection period (Schockaert and Surkyn, 2012). These hypotheses are summarized in 1. Note that we assume that the household formation processes will remain identical during the complete projection horizon. This means that the forecasted population structure (and its impact on inequality) are the result of ageing and the projection of household formation processes of the *current* population. In the case of the educational projection, we assume a slight rise in educational retention.

The Lipro-projection model presents some clear advantages with respect to classical projections by age and sex only. First, the results are much richer. In addition, Lipro is a fully dynamic model where vital events and migration are differentiated by, and interact with, household formation and educational processes. Consequently, important compositional changes mitigate the evolution of fertility, mortality and migration, and impose constraints on future population trends. Moreover, modification in one household position also imposes constraints on the adjustments in other household positions. For example, if for the purpose of population projections we accept the number of same sex couple formation to be negligible, the number of men that transit to the state of "married couple without children "during one projection period should be equal to the number of women entering this state (and vice versa). In other words, Lipro calibrates the theoretically linked transitions. The constraints included in the household projection are explained in Schockaert and Surkyn (2012). For the projection by educational levels, no constraints were used. The above properties of Lipro projections enhance the reliability of the results by ensuring coherence in population trends.

When performing static ageing, we reweigh the *EU-SILC* 2008 data in such a way that the sample characteristics match the forecasted population distribution. For example, the newly calculated weights ensure that the fraction of single mothers aged 20-25 equals the projected fraction for 2011 and 2031. On the basis of these new weights, we now compute the income distributions for 2011 and 2031.

Reweighting by a simple reweighting of cells (Kalton and Flores-Cervantes, 2003) however, is ruled out for two reasons. First, we use two sets of demographic projections - by age, sex and household position and by age, sex and educational attainment. Since these are made independently from each other, they possibly result in conflicting reweighting factors. Although a simple solution seems to be available by calibrating the conflicting individual weights, a second more fundamental problem remains. In the described demographic projections, the unit of observation is the individual, while for our distributional analysis, the unit of observation has to be the household. In order to marry the individual-based demographic projections with the household-based income data, we calibrate the base year sample using household weights, to the individual-based totals implied by the demographic projections (Deville and Sarndal, 1992). Such a calibration procedure finds new weights as close as possible to the old ones, such that the individual-based demographically projected totals are respected. The procedure is explained in detail in De Blander et al. (2013).

Finally, it is important to note that we presume that economic conditions will remain those of the base-year 2008. In the context of the present analysis, we chose to omit potential wage or benefit changes due to future economic growth in order to single out, ceteris paribus, the impact of population change. Nonetheless, labour market participation and wages are estimated on the basis of individual's characteristics. Hereby, we implicitly assume that their correlation with the projected population variables remains constant over time, i.e. as manifested in the observed sample in the base year. Consequently, economic growth related directly to population change is incorporated in the model.

[1] The application of the Lipro-household typology to the Census and register data was discussed intensively by Lesthaeghe et al. (1997) and will be omitted in the current paper.

Period			2006 $-'$ 10	2011 $-'$ 15	2016 $-'$ 20	2021 $-'$ 25	2026 $-'$ 30
Fertility	$\frac{\text{births}}{\text{woman}}$		1.73	1.82	1.76	1.72	1.70
Life expectancy	age (years)	Female	82.7	83.2	83.8	84.4	85.0
		Male	76.9	77.7	78.7	79.6	80.7
External migration	$\frac{\text{rate}}{\text{rate in baseline period}}$		1	1.2	1.4	1.2	1
Educational retention	$\frac{\text{rate}}{\text{rate in baseline period}}$		1	1.065	1.13	1.195	1.25

Table 1: Projection scenarios

2.2 Analysis of inequality and income distribution

2.2.1 Decomposition of inequality and poverty indexes

The Theil index[2] is a inequality measure. It is given by

$$T_F = \int_0^\infty \frac{x}{\mu_F} \ln\left(\frac{x}{\mu_F}\right) dF(x). \tag{1}$$

.It ranges from 0 at complete income equality (everybody earns the same amount) to $\ln N$, when inequality is maximal (only one individual earns total income). It is measured by

$$\hat{T}_F = N^{-1} \sum_{i=1}^{N} \frac{x_i}{\bar{x}} \ln\left(\frac{x_i}{\bar{x}}\right).$$

The main advantage of the Theil index is its sub-group decomposability, i.e.

$$T_F = T_B + \sum_{g=1}^{G} q_g \cdot T_{F_g}, \tag{2}$$

where $g = 1, \ldots, G$ constitutes a partition of the population[3], q_g is the equivalized income *share* of subgroup g, T_{F_g} is the Theil index of sub-group g and T_B is the between group Theil index. It is calculated by attributing every individual the group-mean equivalized income μ_{F_g}.

In the context of analysing the influence of change in the population composition on the income distribution, this property of the Theil is helpful. We use it to analyse the contribution of specific changes in the prevalence of population categories, for example, the increase of single headed households or individuals with higher education, to the forecasted changes in inequality. This contribution can be derived from the second part of the formula $\sum_{g=1}^{G} q_g \cdot T_{F_g}$, since the income share is related to the sub-group population share.

(Foster et al., 1984) introduced the family of poverty indices

$$P^{FGT}(F \mid z, \alpha) = \int_0^\infty \left[\max\left(1 - \frac{x}{z}, 0\right)\right]^\alpha f(x)\, dx,$$

known as the Foster-Greer-Thorbecke class of poverty measures, where, as before, the variable x represents equivalized net household income, $f(x)$ its density function and with z the poverty line or poverty threshold. In this paper, we use a relative poverty line, defined as 60% of the median equivalized income of the *Flemish* population. Note that the poverty line varies over time. $P^{FGT}(F \mid z, 1)$ is the normalized poverty gap[4]. The (non-normalized) poverty gap per poor person is equal to the difference between the individual's income and the poverty line divided by the poverty rate. It is an indicator of the severity of poverty.

A key advantage of the *FGT* class of poverty indices is its subgroup decomposability, i.e.

$$P^{FGT}(F \mid z, \alpha) = \sum_{k=1}^{K} p_k P^{FGT}(F_k \mid z, \alpha),$$

with p_k the fraction of the population belonging to subgroup k and F_k the income distribution within subgroup k. As in the case of the Theil decomposition, we will use the poverty decomposition to gain insight in the impact of specific shifts in the population composition on poverty.

2.2.2 The relative distribution method

To analyse the complete income distribution, we use the relative distribution method proposed by (Handcock and Aldrich, 2002; Handcock and Morris, 1999). Let $F_0(\cdot)$ denote the cumulative distribution function (*CDF*) of an outcome attribute for a reference group and $F(\cdot)$ the *CDF* for the comparison group, assuming both are absolutely continuous with common support. The objective is to study the differences between the distributions of the outcome attribute in the reference and comparison group.

[2]The Theil index is a member of the class of generalized entropy measures.
[3]Under a partition, every individual belongs to *exactly one* subgroup.
[4]The poverty gap is also known as the poverty deficit.

Letting y_0 and y be random samples from $F_0(\cdot)$ and $F(\cdot)$, respectively, the *grade transformation* of Y to Y_0 is defined as the random variable

$$R = F_0(Y),$$

realisations of which are called *relative data*. Its value is the rank that observation y would obtain in the distribution $F_0(\cdot)$. The *CDF* $G(\cdot)$ of R can be expressed as

$$G(r) = F\left(F_0^{(-1)}(r)\right) \qquad 0 \leq r \leq 1,$$

where r represents the proportion of values Y_0 smaller than y, and $F_0^{(-1)}(r) = \inf_y \{y \mid F_0(y) \geq r\}$ is the quantile function of F_0. The probability density function (*PDF*) of R is

$$g(r) = \frac{f\left(F_0^{(-1)}(r)\right)}{f_0\left(F_0^{(-1)}(r)\right)} \qquad 0 \leq r \leq 1.$$

If $F \equiv F_0$, $g(r)$ is uniform and $G(r)$ is a 45° line.

The relative data can be interpreted as the percentile rank that the comparison value y would have in the reference group. The relative *PDF* $g(r)$ can be interpreted as a density ratio: the ratio of the fraction of respondents in the comparison group to the fraction in the reference group at a given level of the outcome attribute $y = F_0^{(-1)}(r)$. Letting the rth quantile of R be denoted by the value y_{r_0} on the original measurement scale, then

$$g(r) = \frac{f(y_{r_0})}{f_0(y_{r_0})} \qquad 0 \leq r \leq 1.$$

The relative *CDF*, $G(r)$, can be interpreted as the proportion of the comparison group whose attribute lies below the rth quantile of the reference group[5].

Let us clarify the use of the relative PDF with a fictitious example

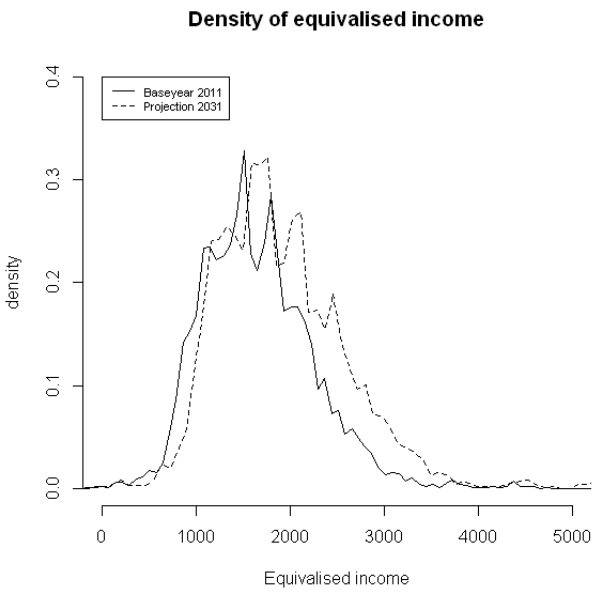
Figure 1 shows the transformation of two income distributions into a relative distribution (density). Sub-figure 1a represents the income distributions in t_1 and t_2; Sub-figure 1b represents the relative income distribution of t_2 compared to t_1 . It shows the ratio of the fraction of individuals at t_2 to the fraction at each income level of the population in t_1 . If the t_1 and t_2 densities would have been exactly the same, with the same median and shape, the relative income distribution would have been a horizontal line at value 1. A value above 1 indicates an increase of the population share of the income level shown on the x-axe; a value below 1 indicates a decrease. In other words, the relative distribution summarizes precisely that part of Sub-figure 1athat we are actually interested in: the space between the curves of t_1 and t_2 . Sub-figure 1bshows that over time the share of the population with an income at the bottom two deciles of the t_1 income distribution will decrease, while higher income levels will show an increase. In the middle of the distribution, shifts are smaller and point to a small decrease of the share of the population around the median.

If the shape of the distribution in t_2 were exactly equal to the shape of the t_1 distribution, a simple comparison of means (or the median) would perfectly describe the income distribution change over time. However, 1indicates that income gains are somewhat larger in the upper than in the lower tale. Handcock and Morris (1999) developed an approach to decompose the overall relative distribution into relative distributions that represent shifts in location and in shape.

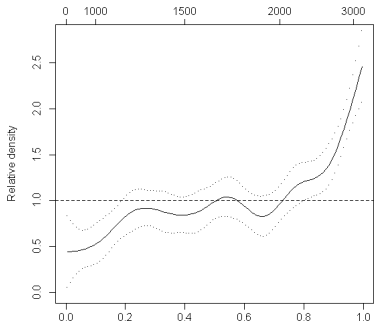
Let Y_A be a hypothetical adjusted population in t_1 that has the median income ("location") of the t_2 population, but its own observed shape. Than $Y_A = Y_0 + \rho$ with $\rho = \text{median}(Y) - \text{median}(Y_0)$. To isolate the location shift, we take the relative distribution of Y_A versus Y_0, in the same way as described in (1). Differences with Y that remain after a location adjustment are differences in "shape". The decomposition of the overall relative density in a shift in location and shape can thus be expressed as

$$g(r) = \frac{f(y_r)}{f_0(y_r)} = \frac{f_A(y_r)}{f_0(y_r)} \cdot \frac{f(y_r)}{f_A(y_r}$$

[5]Note that the relative distribution is scale invariant, i.e. monotone transformations leave it unchanged. For example, one obtains the same relative distribution from a comparison of log-attributes as from the comparison of the attributes.



(a) Densities of equivalized income at time t_1 and t_2



(b) Relative distribution of equivalized income, t_1 versus t_2

Figure 1: Densities and relative distribution of equivalized income, t_1 v t_2

To capture the shape shift in a concise way, we will use the polarization index. The median relative polarization index is the mean absolute deviation around the median of location adjusted relative distribution

$$\mathrm{MRP}\left(F;F_0\right) \quad = \quad 4\mathrm{E}\left[\left|F_0\left(Y-\rho\right)-\frac{1}{2}\right|\right]-1,$$

with $\rho = F^{(-1)}\left(\frac{1}{2}\right) - F_0^{(-1)}\left(\frac{1}{2}\right)$. It varies between -1 and 1, with positive values indicating a proportional shift from more central to less central values. The lower polarization index is the contribution of the lower tail to the *MRP*

$$\mathrm{LRP}\left(F;F_0\right) \quad = \quad 4\mathrm{E}\left[\left|F_0\left(Y-\rho\right)-\frac{1}{2}\right| \mid F_0\left(Y-\rho\right) \leq \frac{1}{2}\right]-1,$$

and similarly for the upper polarization index *URP*.

A decomposition method similar to the median and shape decomposition can be used for co-variate decomposition of the relative distribution. This is done estimating an adjusted population by modifying the t_1 population to have the marginal covariate distribution of the t_2 population and than calculating the income distribution for this adjusted (hypothetical) t_1 population, Y_a. The relative distributions of the adjusted population to the observed t_1 population $F(Y_a)/F(Y_0)$ and to the t_2 population comparison groups $F(Y)/F(Y_a)$ then isolate the compositional and residual effects respectively.

2.3 Quantile regression

The interpretation of the results of a quantile regression is clear

$$q_\alpha \mid x_{it} \quad = \quad \beta_t' x_{it},$$

While *OLS* regression only models the conditional mean, quantile regression models each conditional quantile of the distribution. The advantage of quantile regression is thus that it provides supplementary information about the shape of the distribution. In addition, it also allows for the relationship between covariates and the predicted variable to change over the distribution of the latter. It is this characteristic of quantile regression that we will use to interpret the changing role over time of population characteristics in shaping the income distribution.

3 Population forecasts to 2031

During the second half of the last century, intense modifications in family formation and dissolution were observed (Lesthaeghe et al., 1997): a postponement of first marriage, a reduction of marriage intensity and an increment in divorce rates. These evolutions in combination with the endurance of fertility decline and the increase of life expectancy, will induce a profound change in Flanders´ demographic structure over the next twenty years.

Figure 2 shows the population distribution by age and sex in 2011 and the projection result for 2031. As the baby-boom generations at working and reproductive age in 2011 grow older, the bottom of the pyramid shrinks and the top becomes heavier. Parallel, this ageing process gives rise to changes in the population's household composition as the younger generations of 2011 grow older and their family formation behaviour is reflected in successive age groups in each subsequent projection year. This is demonstrated in Figure 3.

Figure 3 depicts the proportional distribution of household positions by age in 2011 and 2031. On the x-axes are represented the five-year age groups and on the y-axes the proportion of individuals from each age group in single headed households and in couples with and without children. The curves represent the projection years. Panel (a) and panel (b) correspond to women and men respectively. A first important evolution is the decrease of the share of individuals that live in couples with children and a concentration of parenthood around younger ages. These evolutions are related to the descent and postponement of couple formation and to fertility decline leading to a shorter total time span that children are present at their parent's household. Higher divorce rates add to this evolution also to explain the decreasing prevalence of couples without children at more advanced adult age. However, among the elderly population, above the age of 75 for women and 85 for men, the prevalence of individuals living in a couple increases in time. This is easily understood as the result of growing partner's

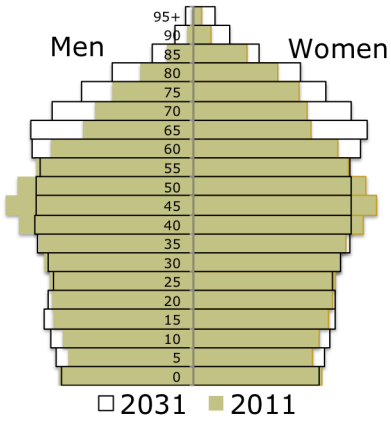
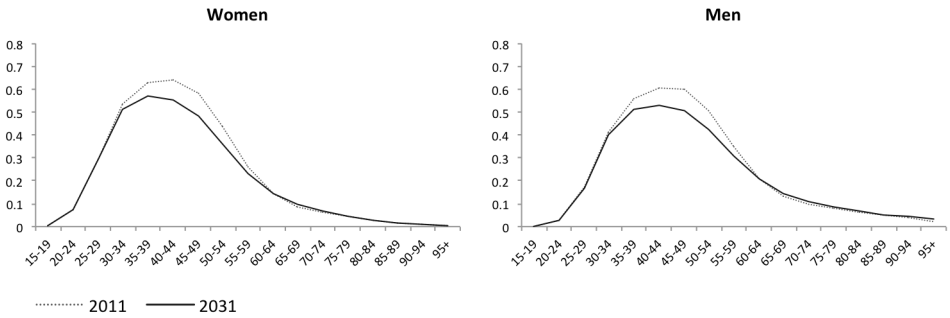
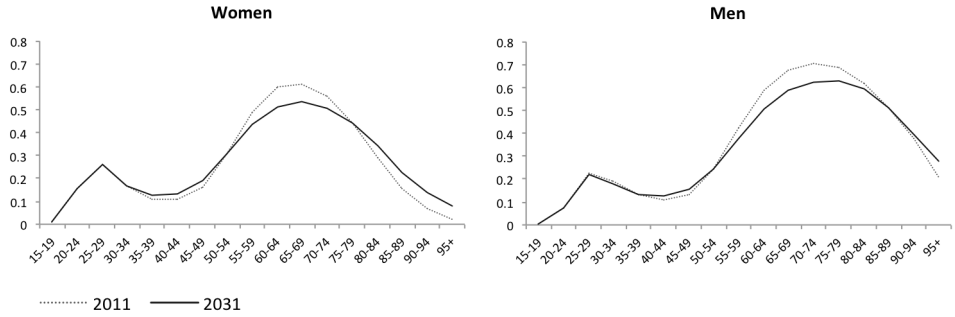


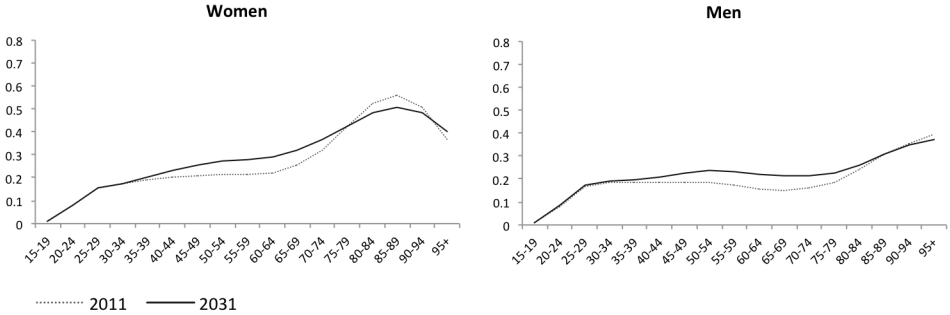
Figure 2: Population pyramid 2011 - 2031



(a) Couples with children



(b) Couples without children



(c) Single headed households

Figure 3: Population by age, sex and Lipro position

survival rates since we projected women's and especially men's life expectancy to grow. The decrease over time of couples results in the rise of single-headed households, most eminent between the age of 40 and 70[6].

Figure 4 depicts the proportional distribution by age of individuals with only primary education or less, lower and higher secondary education and higher education for each five year projection period between 2011 and 2031. Panel (a) and panel (b) represent the distribution for women and men respectively. The Figure shows that parallel to ageing and changes in household position, for all age groups and the two sexes, we foresee a considerable advancement in educational attainment. That is, for each age group the share of the population with higher secondary and higher education is higher in 2031 than in 2011. This process is almost entirely due to the ageing of the population as from each projection year to the next, the educational profile of the younger generations is progressively spread to all ages. Note that within the population under the age of 35, practically no change is observed since we assumed only minor adaptations in educational behaviour (cf. Table 1). Consequently, in the long run, an equalizing tendency between generations and gender is observed. The following comparison demonstrates this tendency: in 2011, the part of the population having at least a higher secondary degree below the age of 40 was about 80% while this was only the case for about 22% of the group above 65. In 2031, 80% of the whole population under 65 will have a higher secondary degree, and this is also the case for almost 60% of the +65 group. Furthermore, in 2011, the younger generations of women had already suppurated the level of the men; in 2031 this will be the case for all age groups, except the most advanced ones

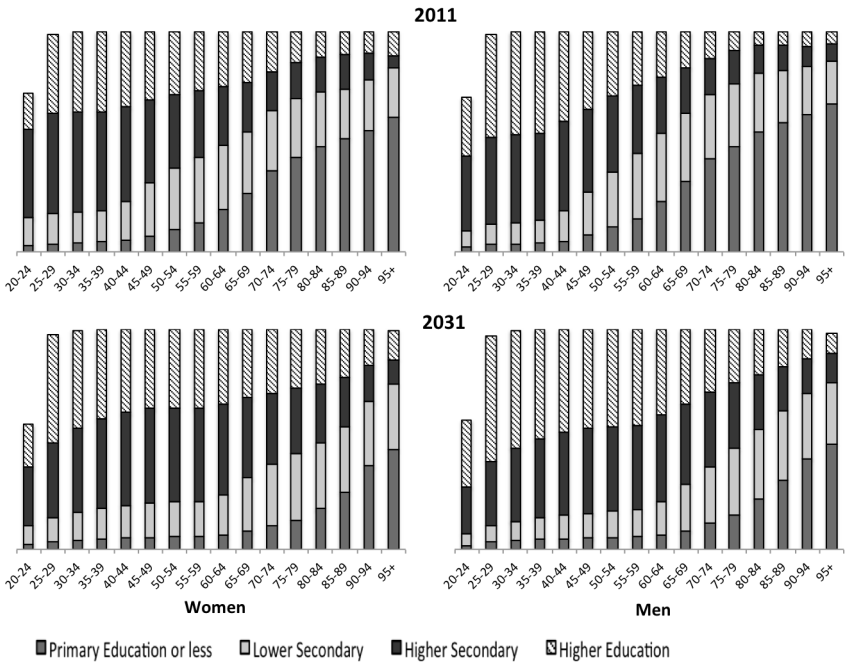
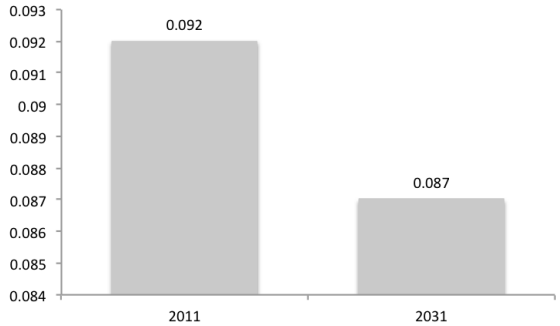


Figure 4: population and educational attainment by age and sex, 2011 and 2031

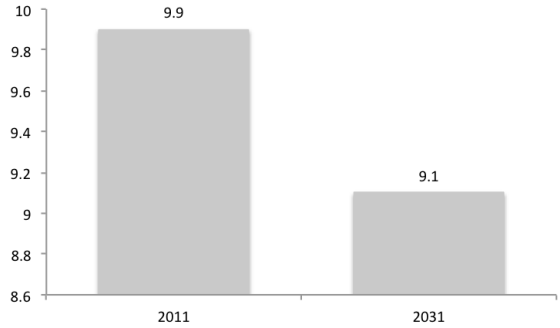
4 The inequality indexes and their decomposition

In this section we depict the forecasted change in income inequality, using the Theil index. In addition, the poverty rate focusses on the evolution of the lower tail of the income distribution. Between 2011 and 2031, both indicators decrease from about 0.092 points and $9,9\%$ to $0,087$ points and $9,1\%$ respectively. In other words, without any changes in economic conditions, between 2011 and 2031, the forecasted population change will induce a decrease in overall inequality and poverty in Flanders. Note however that the change is relatively modest.

[6]Note that, by age and for each projection year, the prevalence of female single headed households progressively outgrows the prevalence of male single headed households. The sex differences in life expectancy and the fact that women on average are younger than their partner, leads to a higher proportion of female than male headed households. At the end of the life course, the proportion of single headed households decreases, as people move to a collective household.



(a) Theil index



(b) Poverty

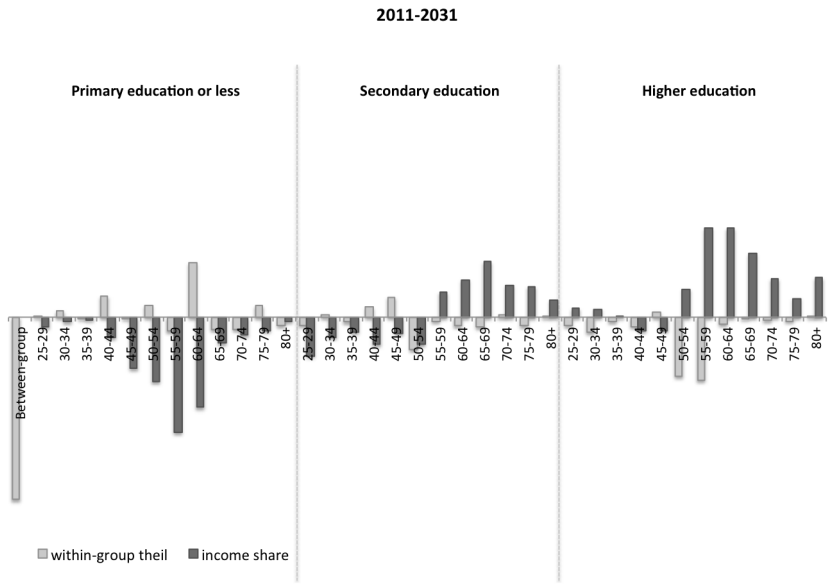
Figure 5: Inequality and poverty indexes, 2031 versus 2011

Figures 6a-6b decompose the population effect on the evolution of the Theil index (cf. Figure 5a). We analyse the contribution of subgroups by age and educational attainment and by age and household position.

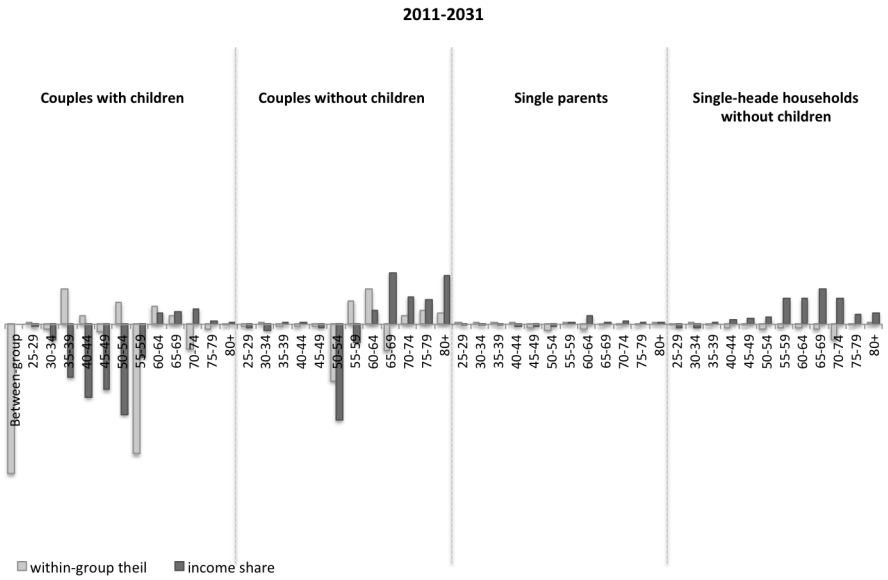
The first bar at the left depicts the change in the between-groups Theil. The other bars indicate each population group's contribution to the overall inequality change. This contribution depends on the *inequality* evolution *within* each group and the group's *share* of the total population's *income,* represented in light and dark grey bars respectively (cf. Equation (2)). The size of the bar indicates the size of contribution to inequality change; upward means a positive impact, downward a negative one. Note that the income share of each group plays the most important role, while within-group shifts have mostly only modest effects. Since we work under the assumption of "no economic change", the within-group inequality change is only due to modifications in the unobserved variable composition (and the sex composition) related to the forecasted population change. Under the same assumption, it is clear that changes in the income share are largely driven by changes in the population; a change in the prevalence of a population category provokes an income share change in the same direction. This way we can directly link the results of the Theil decomposition with the population change described in Section 3.

Sub-figure 6aindicates that the overall decrease in inequality between 2011 and 2031 (Figure 5a) is related to two processes: a reduction of between-group inequality and a decrease in income share of the population with no or only primary education. The latter process is related to the general increase in educational attainment depicted in Figure 4. The reduction of between-group inequality can easily be understood through the attenuation of the generational differences in education. In contrast, Sub-figure 6a also shows that the increase of the income share of individuals over 55 with secondary or higher education, increases inequality. In other words, ageing and the subsequent growth of the elderly population increases inequality, despite their higher educational attainment. This effect is however not enough to offset the impact of the reduction in the lowest educational levels combined with the decrease in between-group inequality.

Figure 6b decomposing the Theil by age and household position also reveals that the growth of the elderly population's income share positively contributes to the Theil, weather living in a couple or in single-headed households. The decrease of the adult population younger than 55 living in couples with children, has a negative impact on inequality. Note however that the within-group inequality among couples under 55 of the



(a) Decomposition of the change in Theil index by age and educational attainment



(b) Decomposition of the change in Theil index by age and household position

Figure 6: Decomposition of the change in Theil index

same age group increases. This can be understood by the larger effect of education growth on the income of double-income households than in households with only one income, leading to increased dispersion. The between-group inequality decreases due to the before mentioned diminishing generational inequality.

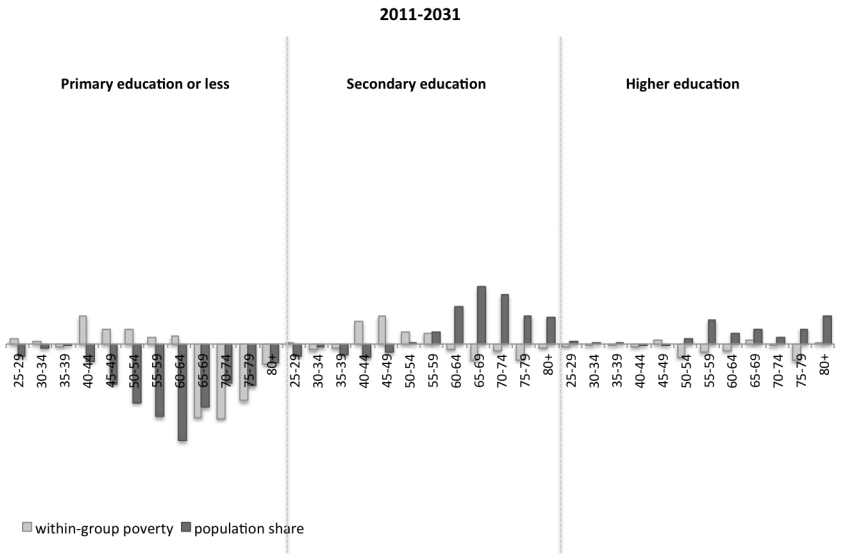


Figure 7: Decomposition of poverty change by age and educational attainment

Figures 7-8 show the decomposition of the declining poverty trend between 2011 and 2031 (cf. Figure 5b). Figure 7 refers to the decomposition by age and educational attainment; Figure 8 to the decomposition by household position. The bars indicate each population group's contribution to the overall poverty change, depending on the within-group poverty and the population share of each group, represented in light and dark grey bars respectively (cf. Equation (2)). The size of the bar indicates the size of contribution to inequality change; upward means a positive impact, downward a negative one.

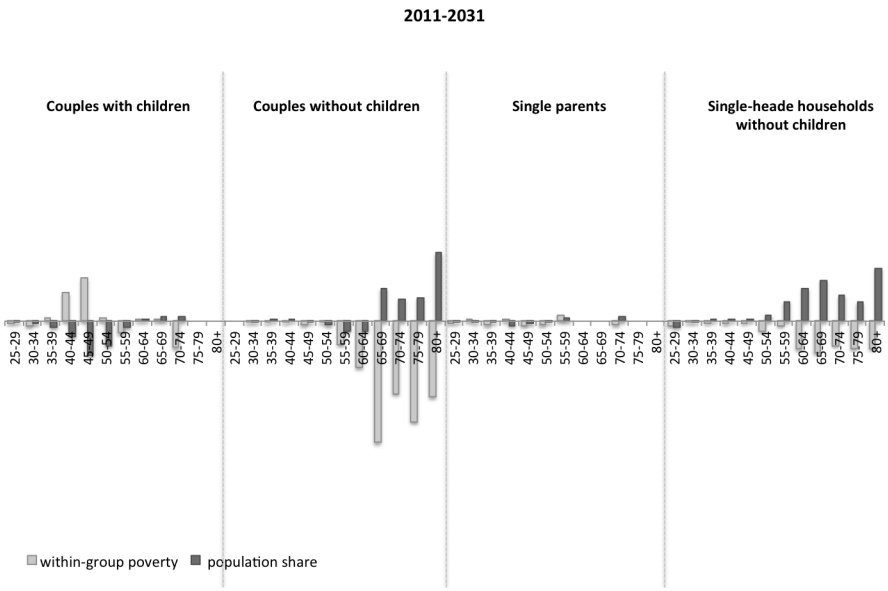


Figure 8: Decomposition of poverty change by age and household position

Figure 7 shows that the decrease of the population share with only or less than primary education contributes to a large extent to the poverty decrease witnessed between 2011 and 2031. Among the lowly educated of more advanced age, the population share change is smaller due to the ageing of the population. We observe however

a prominent reduction in their poverty levels, leading to an important overall contribution to poverty reduction. At higher educational levels (secondary and higher education), ageing increases poverty, despite the decrease of poverty levels within the elderly population. Note that among younger age groups, the within-group poverty increases, except among individuals with higher education. Figure 8 shows that the reduction of poverty within the elderly population is most pronounced among couples without children, but is also witnessed among single headed households. The increase of poverty among the younger population concentrates within couples with children. The decrease of couples without children (cf. Figure 3) counterbalances this tendency, leading to a rather modest joint effect on poverty change. Being single headed households among the lower income groups and couples without children among the higher ones, the increasing share of the former and decreasing share of the latter, also counteracts the negative evolution of the total poverty rate.

5 Analysing the complete income distribution

Section 4 showed that population change between 2011 and 2031 will decrease inequality and poverty. It also indicated what population subgroups contribute to this process, and which ones have a counteracting effect. The analysis however has two mayor shortcomings. First, summary indicators such as the Theil do not tell us what part of the income distribution is actually (most) affected. Is inequality reduced because the lower end of the income distribution caches up, or because the upper end loses ground? Complement the Theil index with the analysis of poverty is only a partial solution. Secondly, a subgroup decomposition falls short on determining the effects of covariates. While we saw for example that ageing increases inequality despite the educational gains of the elderly population, this observation does not say anything about the impact of ageing and educational change *separately*. How much would the impact of ageing have been if no educational gains would have been recorded?

In this section, we use the relative distribution method to complement the analysis of Section 4. Figure 9 represents the income distribution of 2031 relative to 2011. Panel (a) presents the overall relative distribution: the bars indicate the gain (above the horizontal line at value 0) or loss (below the horizontal line at value 0) by the year 2031 of each income decile in 2011 (cf. Subsection 2.2.2). We observe a decrease of the extreme left tale of the distribution, while income concentrates more in the lower middle part of the distribution. The highest income level shares in the population is also emphasised. The graph is is congruent with the observed decrease of the Theil and the poverty rate. They are related to the decrease of the lowest income decile and the growth of the deciles towards the middle of the distribution. Nevertheless, the entropy measure ($entropy = 0.0258$) indicates that the income distribution is only modestly affected by population change.

In Subsection 2.2.2, we explained that the relative distribution is the result of a shift in the income level (median) and of a modifications in the shape of the distribution. This is shown in panel (b) and panel (c) of Figure 9. We observe a slight increase in median income level. The shape modification indicates a larger increase of the highest income decile and the lower middle part of the distribution than one would expect if income increase would have been evenly distributed. The polarization index (cf. subsection) is however not statistically significant.

Figure 10decomposes the effect of the overall population change on the income distribution into the impact of ageing, changes in household composition and educational advancement on the income distribution separately. These pictures show the relative distribution of 2031 versus 2011 if *only* the age, household or educational composition changed, keeping all other marginals as in the baseline year 2008. In absence of any other population change, ageing and changes in household composition increase the lower end of the income distribution. Educational advancement in contrast favours the higher part of the income distribution and decreases the share of the population in at the lower end. In addition, table 2 indicates that educational change increases the polarization in the left tale of the distribtuion, while decreasing polarization in the upper tale. Ageing has the opposite effect.

In other words, the change in prevalence of the different population covariats has counteracting effects on the income distribution. They alter the proportion of individuals with low or high income levels in opposite ways, not only by increasing/decreasing the median income level but also by broading or crushing the distributional shape in opposite directions. The result is the small distributional change and the absence of overall polarization observed in Figure 9.

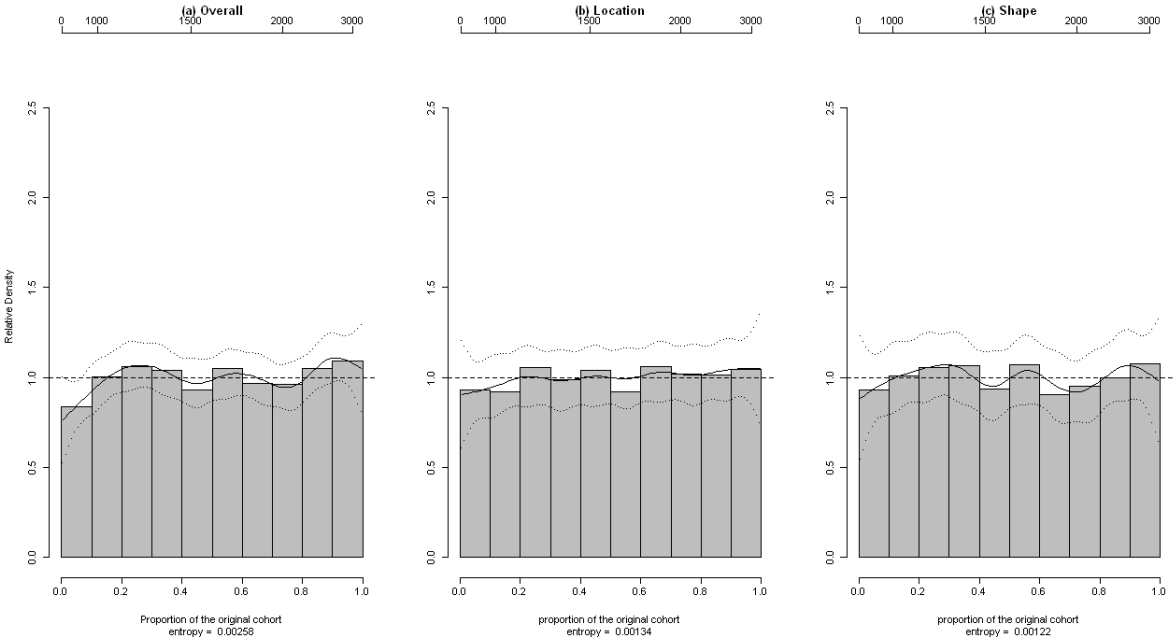
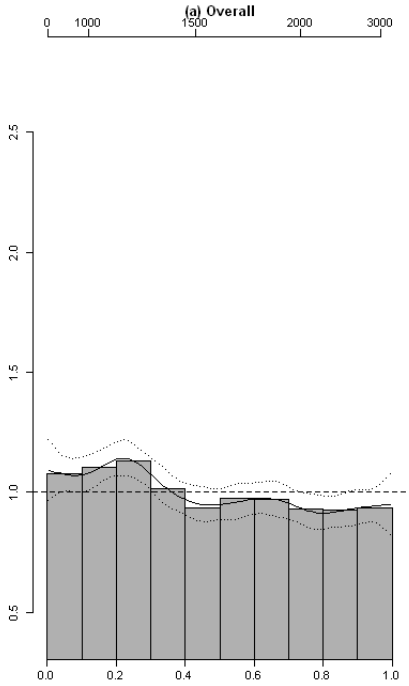


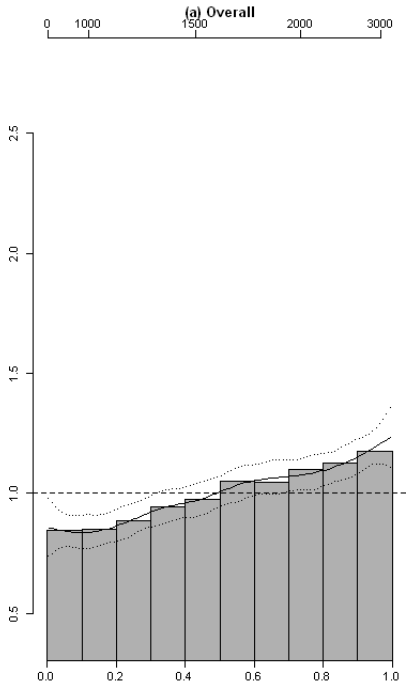
Figure 9: Decomposition of the overall change in income distribution in median and shape shift.

	Ageing		Education		Household comp.	
	index	Pr[|I|>0]	Index	Pr[|I|>0]	Index	Pr[|I|>0]
MPI	-0,008	0,163	0,012	0,055	-0,002	0,409
LPI	-0,042	0,004	0,048	0,001	-0,019	0,113
RPI	0,027	0,044	-0,023	0,061	0,015	0,167

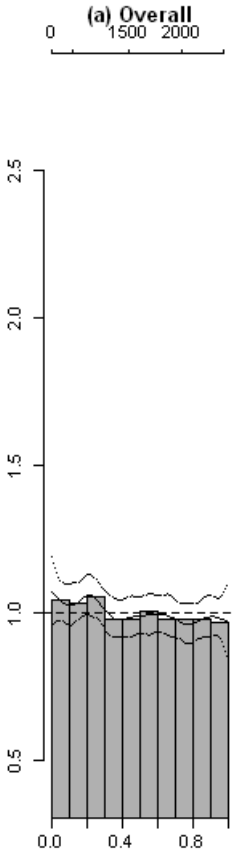
Table 2: Polarization index, age, education and household position



(a) the impact of ageing



(b) The impact of educational growth



(c) The impact of change in household composition

Figure 10: Decomposition of demographic effect in age, household composition and educational effect

6 The changing influence of gender, age, education and household position on household income

Up till know, we measured the impact of change in population composition on the income distribution; that is, the impact of the modification in subgroup or covariate *prevalence.* If we zoom in onto the mechanism of population effects, we can however identify a second complementary influence. After contributing new weights to individuals with different demographic characteristics, at each point in time - 2011 and 2031 - household income is parametric estimation on the basis of the "new" household characteristics. In other words, not only the prevalence, but also the relationship of population covariates with household income changes over time. Figures 11, 12, 13 and 14show the results of a quantile regression of gender, education, household position and age on household income. The first three variables are introduced in the model as dummies using males, less than higher secondary education and couples without children as the reference categories. Age was introduced as a fifth order polynomial. No interaction terms were used. We estimated each income decile, but the results are only shown for the first, third, fifth, seventh and ninth quantile.

Figure 11 shows that both in 2011 and 2031, higher education mainly differentiates household incomes at the upper tail of the distribution. Higher secondary education in contrast increases only modestly from the first to the third decile, differentiating among all income levels fairly alike. Over time, only up to the third decile, the advantage of higher secondary education and higher education is more pronounced in 2031 than in 2011. The impact of higher education even declines at higher income levels[7]. In other words, the educational progression in Flanders shown in Figure 4 attenuates the role of this covariates in determining income differences.

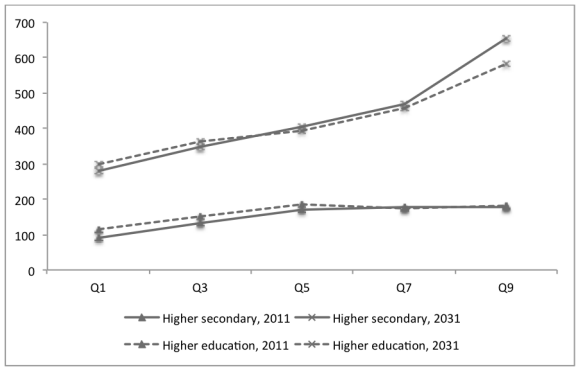


Figure 11: Effect of education on income by quantile, 2031 versus 2011

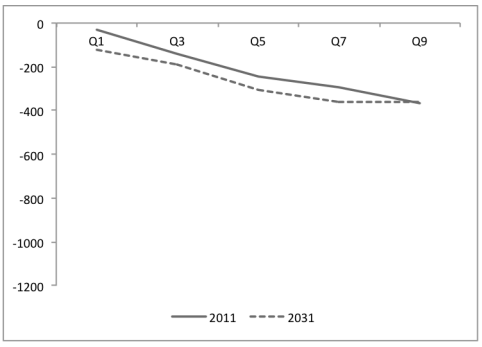
Figure 12 shows the position of couples with children and single headed households with and without children, as compared to couples without children. It is no surprise that the effect of being a single parent is the largest; in 2011, the (equivalized) income difference with couples without children is about €331 in the first decile, and more than €1200 in the ninth decile. Their position remains however stable over time. Most notable is the position of couples with children. As compared to couples without children, in 2011, their household income is about €30 lower in the first decile and more than €365 in the ninth decile[8]. The presence of a child thus has an important impact on the household's income position. The position of single headed households without children is comparable to the one of couples with children, but no noteworthy change is observed in time. Although the position of couples with children shows a tendency to increase over time, for none of the household types change in household income position was statistically significant.

The impact of gender on one's income position is small and only significant in the third and median quantile. Since men and women living in couples by definition have the same household income, the income difference by gender is due to the larger prevalence of single headed households among women.

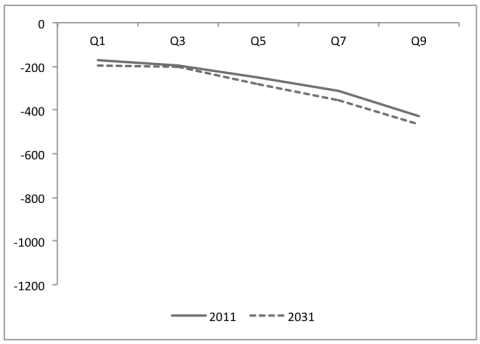
Figure 14 describes the the effect of age on household income. For reasons of simplicity, we only show the results for the first, median and ninth decile. In the first decile, in 2011, individuals between the age of 25 and 50 have an income above the overall median income level. After the age of 50 it drops below the median. In the fifth decile, income differences are larger and the upper limit of the advantaged group moves to the age of 55; in the ninth decile it is over 60. At higher income levels, the disadvantage of the elderly population is thus

[7]The difference is only statistically significant for the ninth decile

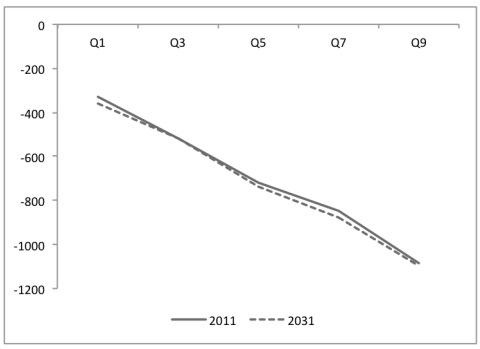
[8]The difference is statistically significant in all deciles but the first.



(a) couples with children



(b) Single headed households without children



(c) Single headed households with children

Figure 12: Effect of household composition on income by quantile, 2031 versus 2011

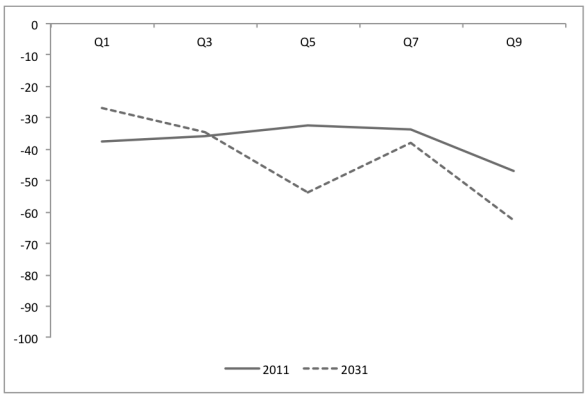


Figure 13: Effect of gender by quantile, 2031 versus 2011

more pronounced and concentrates in higher age categories. Over time, for all three deciles, the position of the elderly improves and the advantage of the adult population diminishes. In other words, the impact of age on income is smaller in 2031 than in 2011.

7 Conclusion

In this paper, we analysed changes in income distribution and inequality between 2011 and 2031 in Flanders, and demonstrated how these are related to population change. We used multi-state population projections (Liproprojections) and reweighting procedures to construct the necessary population and income data in 2011 and 2031. An important feature of our approach is the we assume absence of economic progress; we demonstrate, ex ante, the effect that would have projected population change in absence of all other possible societal or behavioural modifications. We used decomposition of inequality measures, relative distribution methods and quantile regression. These analytical frames provide pieces of the answer to our research question : Can we disentangle the demographic elements most affecting inequality, and single out socio-demographic groups moving to the lower or upper end of the distribution?

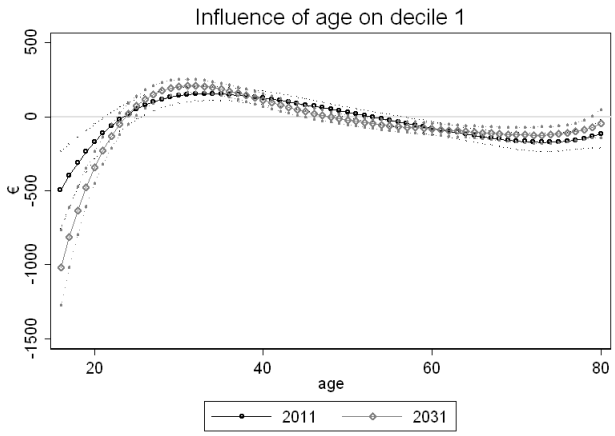
Our population projection models foresee a pronounced ageing of the Flemish population between 2011 and 2031. Parallell, the household and educational composition of the region will also change, as with the ageing process, family formation behaviour and educational levels of each generation in the current population are transferred to successive age groups. Consequently, the projection results demonstrate an increase of single headed households to the detriment of couples, especially with children. Educational attainment levels increase and intergenerational and gender differences in education attenuate.

In the next step of our analysis, we linked these population changes to adjustments in income inequality and poverty. Overall, between 2011 and 2031, we found a modest inequality and poverty reduction. This reduction is related to the decrease of the adult population with no or only primary education as well as the reduction of intergenerational differences in education. We also found an attenuating effect of the decrease in the prevalence of couples with children. The increase of the elderly population inhibits the inequality reduction, independently of the household position and despite their better educational profile.

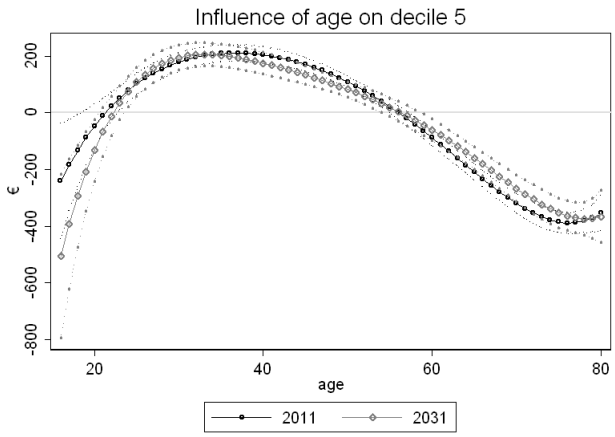
Using the relative distribution method, we analysed the changes in income distribution in more detail. Between 2011 and 2031, we observed a modest increase of the lower middle part of the income distribution and an increase of the highest deciles to the detriment of the sixth and eighth. Nevertheless, the overall modification in income distribution due to population change is rather modest, as the Theil and poverty rate already suggested. A counter-factual decomposition of the relative distribution revealed counteracting influences of the different aspects of population change, clarifying this overall modest transformation of the income distribution. Educational advancement increases the share of people with higher income levels and decreases the lower end of the distribution; ageing and modifications in the population's household composition have exactly the opposite effect. In addition, education narrows the left tale of the distribution, while increases polarization of the upper tale. Ageing does exactly the opposite.

Moreover, we showed that the covariate-income relationship changed over time. Using quantile regression, we indicated that educational differences in income are higher in 2031 than in 2011 among the population with an income below the median income, but lower at higher income levels. Generational differences decrease in time. These changes in covariate-income relationship counteract to the effects of compostional changes on inequality; the impact of educational enhancement would have been larger if no attenuation of the education-income relationship would have taken place among the high income levels. The same holds for the ageing process: the decrease in generational income differences inhibits the increase of the lower end of the distribution due to ageing.

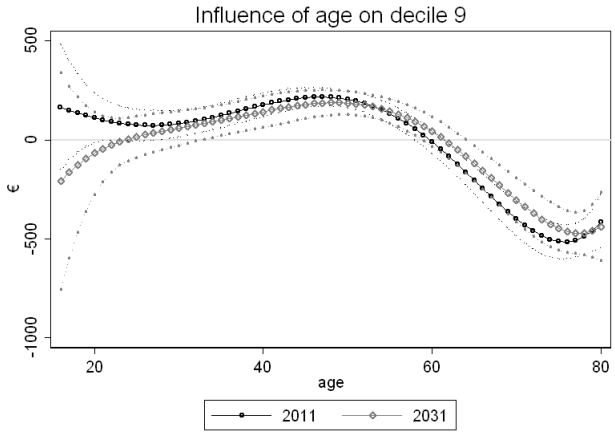
In short, the relation between population change and income distribution involves different components that often act in different directions. Ageing, as expected, does have an negative impact on the income distribution and inequality, but the process comes with at least two covarying population characteristics - household composition and education - that may well have other effects on income. In addition, our paper showed that the analysis of income distribution and population should go beyond the examination of the mean or the median, and should involve looking at compositional effects as well as the change in covariate-income relationship.



(a) First decile



(b) Median



(c) Ninth decile

Figure 14: Effect of age by quantile, 2031 versus 2011

References

Blinder, A. S., 1973. Wage discrimination: Reduced form and structural estimates. Journal of Human Resources 8 (4), 436–455.

Bourguignon, F., Spadaro, A., 2006. Microsimulation as a tool for evaluating redistribution policies. Journal of Economic Inequality 4 (1), 77–106.

Breen, R., Salazar, L., 2010. Has increased women's educational attainment led to greater earnings inequality in the United Kingdom? A multivariate decomposition analysis. European Sociological Review 26, 143–157.

Burtless, G., 1999. Effects of growing wage disparities and changing family composition on the U.S. income distribution. European Economic Review 43 (4-6), 853–865.

Daly, M. C., Valletta, R. G., 2006. Inequlity and poverty in the United States: The effects of rising dispersion and men's earnings and changing family behaviour. Economica 73 (289), 75–98.

De Blander, R., Schockaert, I., Decoster, A., Deboosere, P., 2013. The impact of demographic change on policy indicators and reforms. Flemosi Discussion Paper DP25, CES - KU Leuven.

Deville, J.-C., Sarndal, C.-E., 1992. Calibration estimators in survey sampling. Journal of the American Statistical Association 87 (418), 376–382.

DiNardo, J., Fortin, N. M., Lemieux, T., 1996. Labor market institutions and the distribution of wages, 1973-1992: A semiparametric approach. Econometrica 64 (5), 1001–1044.

Esping-Andersen, G., 2007. Sociological explanations of changing income distributions. Americal Behavioral Scientist 50 (5), 639–658.

Foster, J., Greer, J., Thorbecke, E., 1984. A class of decomposable poverty measures. Econometrica 52 (3), 761–766.

Handcock, M. S., Aldrich, E. M., 2002. Applying relative distribution methods in R. Working Paper 27, Center for Statistics and the Social Sciences, University of Washington.

Handcock, M. S., Morris, M., 1998. Relative distribution methods. Sociological Methodology 28 (1), 53–97.

Handcock, M. S., Morris, M., 1999. Relative Distribution Methods in the Social Sciences. Statistics for Social Science and Behavorial Sciences. Springer.

Hyslop, D. R., Maré, D. C., 2005. Understanding New Zealand's changing income distribution, 1983-1998: A semi-parametric analysis. Economica 72 (287), 469–495.

Kalton, G., Flores-Cervantes, I., 2003. Weighting methods. Journal of Official Statistics 19 (2), 81–97.

Kollmeyer, C., 2012. Family structure, female employment, and national income inequality: A cross-national study of 16 Western countries. European Sociological Review ?, ??–??
URL `http://esr.oxfordjournals.org/content/early/2012/07/03/esr.jcs060.abstract`

Lesthaeghe, R., Deboosere, P., Surkyn, J., Boulanger, P. M., Lambert, A., 1997. Huishoudens en Gezinnen in België. Census Monograph No.4. National Institute for Statistics Belgium, Brussels.

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

Oaxaca, R., 1973. Male-female wage differentials in urban labor markets. International Economic Review 14 (3), 693–709.

Peichl, A., Pestel, N., Schneider, H., 2012. Does size matter? the impact of changes in household structure on income distribution in Germany. Review of Income & Wealth 58 (1), 118–141.

Schockaert, I., Surkyn, J., 2012. Snapshots of our future families. Flemosi Discussion Paper 12, VUB.

Studiedienst Vlaamse Regering, 2011. Svr-projecties van de bevolking en de huishoudens voor vlaamse steden en gemeenten, 2009–2030. SVR-internet page, Brussels.

Van Imhoff, E., Keilman, N. W., 1991. LIPRO 2.0: An Application of a Dynamic Demographic Projection Model to Household Structure in the Netherlands. Swets and Zeitlinger B.V., Amsterdam.