Review of Income and Wealth
Series 56, Number 1, March 2010

# ON MEASURING AND EXPLAINING SOCIOECONOMIC POLARIZATION IN HEALTH WITH AN APPLICATION TO FRENCH DATA

by Bénédicte H. Apouey*

*University of South Florida*

This paper proposes two original measures of socioeconomic polarization, in order to quantify phenomena that are not always taken into account by social inequality measures. Our approach is inspired by the literature on bivariate inequality (the concentration index) and univariate polarization. Like the concentration index, our social polarization measures can be easily computed thanks to a "convenient" regression and decomposed into their determinants. Moreover changes in polarization can also be decomposed into their causes. The paper also provides an empirical illustration of our methods for the probability of reporting excellent or very good health, using cross-sectional data on French women. The findings suggest that after 65 years of age, social polarization in this probability decreases whereas social inequality remains stable. Consequently social polarization conveys additional information to that contained in the concentration index.

## 1. Introduction

Reducing health inequalities between the poor and the better-off and analyzing their causes are two objectives of health policies. This explains why the measurement of income-related dispersion in health and the decomposition of measures into factors are crucial.

The standard measurement tool is a bivariate measure of inequality, the concentration index, which was first introduced by Wagstaff *et al.* (1989). It measures the difference between the observed distribution of health and income in which health and income are positively correlated, and a hypothetical distribution in which an individual's health and his income are independent. This index has now become a standard measurement tool in the economic literature on inequality in health.

However empirical analyses may suggest that this inequality tool does not always give an adequate picture of polarization phenomena. For example, the bivariate density of health—more precisely the probability of reporting excellent or very good health—and income suggests a polarization phenomenon for French women between 45 and 64 years of age, followed by a depolarization phenomenon after 65. This means that we observe first the formation of two distinct peaks, one

*Note*: Data were supplied by the INSEE and the IRDES. I thank two anonymous referees, Andrei Barbos, Valérie Bérenger, Andrew Clark, Lionel Da Cruz, Joseph Deutsch, Brigitte Dormont, Marc Fleurbaey, my advisor Pierre-Yves Geoffard, Jacques Silber, Eddy Van Doorslaer, Shlomo Yitzhaki and participants at the Research Workshop of the Israel Science Foundation on "Income Polarization: Measurement, Determinants and Implications" (May 2008) for useful comments and suggestions.

*Correspondence to: Bénédicte H. Apouey, University of South Florida, Department of Economics, 4202 East Fowler Avenue, BSN 3403, Tampa, FL 33620-5500, USA (benedicte.apouey@gmail.com).

being made of the wealthier and healthier, and the second one consisting of the poorer and less healthy, and then the concentration of the distribution in one single pole. The point is that the traditional inequality measure does not decrease when depolarization occurs. This finding demonstrates that some aspects of variables' dispersion are not conveyed by inequality measures and that polarization measures may be a good complement.

As a consequence in this paper we highlight the need for a socioeconomic polarization measure in the health field using a sample of French women, and develop an original tool to quantify social polarization. We also show that our polarization measures can be easily computed using a "convenient" regression and decomposed into factors, which enables us to explain polarization.

The paper is organized as follows. Section 2 summarizes the measurement methodology on inequalities in the health literature. Section 3 contains our main theoretical results, since we propose our indices of social polarization, show that they can be easily computed, and develop a decomposition to unravel the causes of polarization. Section 4 contains the empirical illustration: we highlight the usefulness of our measures when the inequality measure fails and then analyze the evolution of social polarization with age for French women. Section 5 concludes.

## 2. Health Inequalities

This section contains a summary of the inequality tools used in the health field.

Our intention is first to highlight a way to derive a social dispersion measure from an overall dispersion index. Indeed in the literature on health inequalities two distinct strands are evident (Wagstaff and Van Doorslaer, 2004). The first one is a univariate setting that analyzes overall or pure health inequalities. In this approach all inequalities in health are measured, irrespective of the social characteristics of the individuals. This first approach is quite rare in the literature since most researchers consider that it is not dispersion in health which is interesting, but income-related dispersion in health. The second strand is a bivariate or social approach, in which the relationship between socioeconomic characteristics (such as income) and health is taken into account. As the reader will see in what follows, the difference between the univariate and the bivariate inequality measures is that for the univariate measure individuals are ranked by health levels, beginning with the least healthy, whereas for the bivariate measure individuals are ranked by income, beginning with the poorest. Section 3 makes clear that we also use this simple change in the ranking of individuals to derive our bivariate polarization measure from a well-known univariate polarization index.

The second goal of this section is to present the properties of the bivariate inequality measure because we took inspiration from them to develop a polarization index with "good" properties. Indeed the bivariate inequality measure can be easily computed thanks to a "convenient" regression and decomposed into the inequalities' causes. As shown in Section 3, our social polarization measures satisfy very similar properties.

In the rest of the paper, we assume without loss of generality that there are $n$ individuals and that $n$ is odd.

### 2.1. *From Overall to Social Inequalities for Health*

In the univariate approach, we assume that individuals are ordered by increasing health. So if we denote $y_i$ the positive health of the $i$-th individual, then the vector of health is $y = (y_1, \ldots, y_n)$ with $y_1 \leq \ldots \leq y_n$. The traditional measure of overall health inequalities is the Gini index for health (Wagstaff *et al.*, 1991) which is equal to:

$$G(y): \Re_+^n \rightarrow [0,1]$$

$$G(y) = \frac{2}{n\bar{y}} \sum_i y_i R_{y,i} - 1$$

with $\bar{y}$ the mean health and $R_{y,i}$ the relative rank of individual $i$: $R_{y,i} = \frac{2i-1}{2n}$.

In the bivariate approach, income-related health inequalities are measured using the concentration index for health. Here for the calculation of the relative rank, individuals are ranked according to income unlike in the univariate setting: rank 1 is given to the poorest individual and rank $n$ to the wealthiest individual. So if we denote $w_i$ the positive income of the $i$-th individual, then the vector of health statuses and incomes are $y = (y_1, \ldots, y_n)$ and $w = (w_1, \ldots, w_n)$ with $w_1 \leq \ldots \leq w_n$. The concentration index is defined as:

$$C(y,w): \Re_+^{2n} \rightarrow [-1,1]$$

$$C(y,w) = \frac{2}{n\bar{y}} \sum_{i=1}^{n} y_i R_{w,i} - 1$$

where $R_{w,i}$ is the relative rank of individual $i$, $R_{w,i} = \frac{2i-1}{2n}$.

$C(y, w)$ lies between −1 and 1. It decreases as health inequalities favor more the poorer members of society, and increases as health inequalities favor more the richer individuals.

### 2.2. *Computation of Social Health Inequalities*

The concentration index can be estimated thanks to a regression (Kakwani *et al.*, 1997):

$$2\sigma_{R_w}^2 \frac{y_i}{\bar{y}} = \alpha + \gamma R_{w,i} + u_i$$

where $\sigma_{R_w}^2$ is the variance of the relative rank and $\alpha$ is the intercept term.

Simple calculations show that the estimator $\hat{\gamma}$ is equal to the concentration index:

$$C(y,w) = \hat{\gamma}.$$

### 2.3. *Decomposition of Social Health Inequalities*

In order to understand the causes of social health inequalities, the concentration index can be decomposed into factors (Wagstaff *et al.*, 2003). Thus if health is explained by a number of factors $x_k$ using an additive function:

$$y_i = \alpha + \sum_k \beta_{x_k} x_{ki} + \varepsilon_i$$

then decomposition is given by:

$$C(y,w) = \sum_k \frac{\beta_{x_k}\bar{x}_k}{\bar{y}} C(x_k,w) + \frac{GC_\varepsilon}{\bar{y}} \tag{1}$$

where $\bar{x}_k$ denotes the mean of $x_k$, $C(x_k, w)$ is the concentration index for $x_k$, and $GC$ is a generalized concentration index for $\varepsilon_i$, $GC_\varepsilon = \frac{2}{n}\sum_{i=1}^{n} \varepsilon_i R_{w,i}$.

Equation (1) underlines that social inequalities are due to two components. The first component $\sum_k \frac{\beta_{x_k}\bar{x}_k}{\bar{y}} C(x_k,w)$ is an explained component. It is equal to a weighted sum of concentration indices of factors $x_k$, where the weights are the elasticities of $y$ with respect to $x_k$ evaluated at the mean $(\bar{y}, \bar{x}_k)$. Moreover the presence of $C(x_k, w)$ in the decomposition indicates that a factor $x_k$ has an effect on health distribution across income levels only if it is itself unequally distributed across income levels. The second component of equation (1), $\frac{GC_\varepsilon}{\bar{y}}$, is an unexplained component or a residual.

This decomposition method has become very common in health economics (Van Doorslaer and Jones, 2003; Wagstaff *et al.*, 2003; Van Doorslaer and Koolman, 2004).

## 3. Health Polarization

Up to now dispersion in health has been measured using inequality tools only. However some aspects of dispersion are not captured by inequality measures. Indeed assume that a sample of individuals is made up of two groups: first, individuals whose income is smaller than the median income and whose health is poor; and second, individuals whose income is greater than or equal to the median income and whose health is good. Local equalizations of health differences among the two groups will decrease social inequality,[1] but they will also lead to two better defined groups and so these equalizations will increase polarization between the groups. This difference between the evolution of inequality and polarization highlights that inequality and polarization are distinct concepts.

This theoretical argument in favor of the use of polarization measures to complement inequality measures is reinforced by mere observation of data. Indeed

[1]When the social inequality measure satisfies a transfer principle.

they sometimes highlight the emergence of two poles in distributions (see Section 4). Such phenomena are more naturally handled with polarization tools than with inequality measures.

This section thus presents our social polarization measures and contains the main results of the paper.

### 3.1. *Overall Health Polarization*

Overall polarization can be quantified using a number of measures (see Deutsch *et al.*, 2007). Among them we concentrate on Wolfson's index (1994) and we apply it to the health vector. We thus assume that individuals are ranked according to their health level, and that the population is divided into two groups by the median health value. Then the index is defined as:

$$P(y):\Re_+^n \rightarrow [0,1]$$

$$P(y)=\frac{2\bar{y}}{m(y)}[2(0.5-L(0.5))-G(y)] \qquad (2)$$

where $m(y)$ denotes the median health and $L(0.5)$ is the value of the Lorenz curve for health at the 50th population percentile.

The index lies between 0 and 1. When the index is equal to zero then all individuals have the same health level. This absence of polarization coincides with perfect equality. When polarization is maximum, $P = 1$, then one half of the population has the lowest health ($y = 0$) whereas half of the population has excellent health.

Rodríguez and Salas (2003) give an interesting interpretation of this index:

$$P(y)=\frac{2\bar{y}}{m(y)}\left(G^B(y)-G^W(y)\right)$$

where $G^B$ and $G^W$ are respectively the between-groups and the within-groups Gini indices.

So $P$ is proportional to the difference between the between-group Gini and the within-group Gini. As a consequence a decrease of inequalities among the groups implies a rise of polarization. In other words $P$ satisfies the "increased bipolarity axiom" developed by Wang and Tsui (2000). Rodríguez and Salas (2003) also show that it satisfies the "increased spread axiom."

### 3.2. *The Social Polarization Measure for Health*

The difference between the traditional overall and social inequality measures used in the health field is the ranking of individuals. We thus propose to create our bivariate polarization measures by re-ordering individuals by income (instead of health). Since we are interested in polarization, we also assume that there are two groups (or poles):

- The first group contains individuals whose income is strictly smaller than the median income. So the ranks of these individuals go from 1 to $\frac{n-1}{2}$.
- The second group is made of individuals whose income is greater than or equal to the median income. In this group the ranks of individuals go from $\frac{n+1}{2}$ to $n$.

We then quantify income-related health polarization using Wolfson's index and a modified Wolfson's index for this new distribution. Our indices of social health polarization are defined as:

**Definition 1.**

$$SP_1(y,w):\Re_+^{2n}\to\Re$$

(3) $$SP_1(y,w)=\frac{2\overline{y}}{m_y}\left[2\left(0.5-L_{y,w}(0.5)\right)-C(y,w)\right]$$

$$SP_2(y,w):\Re_+^{2n}\to[-1,1]$$

(4) $$SP_2(y,w)=2\left(0.5-L_{y,w}(0.5)\right)-C(y,w)$$

*where $m_y$ is the health of the individual with the median income and $L_{y,w}(0.5)$ is the concentration curve for health evaluated at the 50th percentile.*

$SP_1$ and $SP_2$ are minimum ($SP_1=-\frac{2\overline{y}}{m_y}$ and $SP_2=-1$) when the health of the individuals from the richer group is poor ($y_i=0$ for $i=\frac{n+1}{2},\ldots,n$), whereas the health of the individuals from the poorer group is excellent. On the contrary polarization is maximum ($SP_1=\frac{2\overline{y}}{m_y}$ and $SP_2=1$) when the health of the individuals from the poorer group is poor ($y_i=0$ for $i=1,\ldots,\frac{n-1}{2}$), whereas that of the individuals from the richer group is excellent.

The use of Wolfson's index in a bivariate perspective is all the more pertinent as the indices we get have an intuitive interpretation and satisfy a number of interesting properties. First, social polarization indices can be interpreted as a function of the difference between social inequality between the groups and social inequality within the groups. To highlight this, we need to introduce a new notation.

Let $y^M$ be a modified health vector such that in the two groups individual health is replaced by the group mean health:

$$y_i^M=\begin{cases}\frac{2}{n-1}\sum_{i=1}^{(n-1)/2}y_i & \text{for } i=1,\ldots,\frac{n-1}{2}\\ \frac{2}{n+1}\sum_{i=(n+1)/2}^{n}y_i & \text{for } i=\frac{n+1}{2},\ldots,n.\end{cases}$$

The between-groups concentration index $C^B$ is defined as the concentration index obtained if individual health $y_i$ in the two groups were to be replaced by the group mean health.

**Definition 2.** $C^B$.
*The between-groups concentration index is:*

$$C^B(y,w) = C(y^M, w)$$

(5)
$$C^B(y,w) = \frac{2}{n\bar{y}}\sum_{i=1}^{n} y_i^M R_{w,i} - 1.$$

As for the within-groups concentration index $C^W$, this is defined as the weighted sum of the concentration indices within the groups (where the weights are the share of health in each group), and this is equal to the difference between the concentration index and the between-groups concentration index.

**Definition 3.** $C^W$.
*The within-groups concentration index is:*

$$C^W(y,w) = C - C^B(y,w)$$

$$C^W(y,w) = \frac{2}{n\bar{y}}\sum (y_i - y_i^M) R_{w,i}.$$

The intuition of Rodríguez and Salas (2003) still applies so the indices can be reformulated as:

(6)
$$SP_1(y,w) = \frac{2\bar{y}}{m_y}\left(C^B(y,w) - C^W(y,w)\right)$$

(7)
$$SP_1(y,w) = \frac{2\bar{y}}{m_y}\left[\frac{2}{n\bar{y}}\sum_{i=1}^{n}(2y_i^M - y_i)R_{w,i} - 1\right]$$

(8)
$$SP_2(y,w) = C^B(y,w) - C^W(y,w)$$

(9)
$$SP_2(y,w) = \frac{2}{n\bar{y}}\sum_{i=1}^{n}(2y_i^M - y_i)R_{w,i} - 1.$$

These reformulations enable us to show that $SP_1$ and $SP_2$ satisfy two axioms that are similar in a bivariate context to the "increased bipolarity axiom" and the "spread axiom" developed by Wang and Tsui (2000) in a univariate context. To show that we need to define an analog to the univariate transfer principle in a bivariate context.

**Definition 4.** *Equalizing transfer in a bivariate perspective.*
*Let i and j denote the ranks of two individuals with $i < j$ such that $w_i < w_j$. A distribution $(y^2, w)$ is derived from a distribution $(y^1, w)$ by an equalizing transfer if $y_i^2 = max(y_i^1, y_j^1)$, $y_j^2 = min(y_i^1, y_j^1)$, $y_l^2 = y_l^2$ for $l \neq i, j$. Such an equalizing transfer is strict if $y^2 \neq y^1$.*

The strict equalizing transfer is a health transfer from individual $j$ to individual $i$ when $j$ has both better health and higher income than $i$.[2] A strict equalizing transfer implies that the concentration curve for health vector $y^2$ is above that for $y^1$.

The equalizing transfer is clearly inspired by the literature on correlation increasing transfers (see Atkinson and Bourguignon, 1982; Boland and Proschan, 1988; Tsui, 1999; Abul Naga and Geoffard, 2006). However there are two minor differences:

- A correlation increasing transfer increases inequality, whereas our equalizing transfer decreases inequality.
- In a bivariate perspective, a correlation increasing transfer is either a health transfer or an income transfer, whereas our equalizing transfer is only a health transfer. In fact we do not need to define an income transfer between two individuals in our approach.[3]

This definition of an equalizing transfer permits us to develop two axioms: increased spread and increased bipolarity in a bivariate context.

**Axiom.** *Increased spread (in a bivariate perspective).*
*For any $(y^1, w)$, $(y^2, w)$, polarization for $(y^1, w)$ is strictly greater than polarization for $(y^2, w)$ if $(y^2, w)$ is derived from $(y^1, w)$ via a sequence of strict equalizing transfers between individuals i and j such that $i \leq \frac{n-1}{2}$ and $j \geq \frac{n+3}{2}$.*

This axiom means that polarization decreases when there is an equalizing transfer between the two groups, more precisely from an individual whose income is strictly greater than the median income to an individual whose income is strictly smaller than the median income.

Let $y^-$ be the subvector including $y_i$ such that $i < \frac{n+1}{2}$ and $y^+$ the subvector including $y_i$ such that $i > \frac{n+1}{2}$. Similarly let $w^-$ be the subvector including $w_i$ such that $i < \frac{n+1}{2}$ and $w^+$ the subvector including $w_i$ such that $i > \frac{n+1}{2}$. Since $m_y$ is the

[2]There is no denying that a health transfer is practically impossible, however the concept of health transfer will help us to compare two distributions in terms of polarization levels. So we will do it "as if" a distribution was derived from another one thanks to a health transfer.

[3]Indeed assume that $p$ and $q$ denote the ranks of two individuals such that $y_p < y_q$. Then a distribution $(y, w^4)$ is derived from a distribution $(y, w^3)$ by a transfer if $w_p^4 = max(w_p^3, w_q^3)$, $w_q^4 = min(w_p^3, w_q^3)$ and $w_l^4 = w_l^3$ for $l \neq p, q$. When $w^3 \neq w^4$ this transfer is a strict transfer of income from individual $q$ to individual $p$.

The point is that $w^4$ could also have been obtained from $w^3$ by an equalizing transfer of health from the individual who has the larger income $max(w_p^3, w_q^3)$ to the individual who has the lower income $min(w_p^3, w_q^3)$. So there is no need to define an equalizing transfer of income in our approach.

health of the individual whose income is the median income and $m(w)$ is the median income, then $y = (y^-, m_y, y^+)$ and $w = (w^-, m(w), w^+)$.

**Axiom.** *Increased bipolarity (in a bivariate perspective).*
*For any $(y^1, w)$, $(y^2, w)$, polarization for $(y^2, w)$ is strictly greater than polarization for $(y^1, w)$ if $(y^{2-}, w^-)$ is derived from $(y^{1-}, w^-)$ via a sequence of strict equalizing transfers or $(y^{2+}, w^+)$ is derived from $(y^{1+}, w^+)$ via a sequence of equalizing transfers.*

**Result 1.** *$SP_1$ and $SP_2$ satisfy the Increased spread and Increased bipolarity axioms.*

**Proof.** *See Appendix, Proof 1.*

Computation of the Social Polarization Measures

Our social polarization indices are all the more interesting as they can be easily computed using econometric techniques.

**Result 2.** *The social polarization indices can be estimated thanks to a regression:*

$$2\sigma_{R_w}^2 \frac{2y_i^M - y_i}{\bar{y}} = \alpha + \gamma R_{w,i} + u_i \tag{10}$$

*where $\sigma_{R_w}^2$ is the variance of the relative rank, and $\alpha$ is the intercept term.*

*Simple calculations show that:*

$$SP_1(y, w) = \frac{2\bar{y}}{m_y}\hat{\gamma}$$
$$SP_2(y, w) = \hat{\gamma}.$$

**Proof.** *See Appendix, Proof 2.*

Compared to the direct computation of the indices, this technique permits us to obtain a standard error for $SP_2$.[4] This standard error is not wholly accurate because of serial correlation in the errors as a result of the presence of the relative rank variable. However for the social polarization measures, taking into account serial correlation and sample design features make little difference, like for the concentration index (O'Donnell *et al.*, 2006).

Decomposition of Social Health Polarization

We now show that like the concentration index, our social polarization measures can be decomposed into their causes.

[4]Standard error for $SP_1$ can be computed using several techniques. For example, the Stata nlcom command (http://www.stata.com/help.cgi?nlcom) computes it using the delta method.

**Result 3.** *For a linear regression model,*

$$y_i = \alpha + \sum_k \beta_{x_k} x_{ki} + \varepsilon_i \tag{11}$$

*and for n odd and large, the social polarization index for y can be approximated by,*

$$SP_1(y,w) \approx \overbrace{\underbrace{\sum_k \underbrace{\frac{\beta_{x_k} m_{x_k}}{m_y}}_{\text{Elasticity}} SP_1(x_k, y)}_{\text{Contribution of } x_k}}^{\text{Explained}} + \overbrace{\frac{2\bar{y}}{m_y} GSP_\varepsilon}^{\text{Unexplained}} \tag{12}$$

$$SP_2(y,w) \approx \overbrace{\underbrace{\sum_k \underbrace{\frac{\beta_{x_k} \bar{x}_k}{\bar{y}}}_{\text{Elasticity}} SP_2(x_k, w)}_{\text{Contribution of } x_k}}^{\text{Explained}} + \overbrace{GSP_\varepsilon}^{\text{Unexplained}} \tag{13}$$

*where* $m_{x_k}$ *is the level of* $x_k$ *for the individual with the median income,* $SP_1(x_k, w)$ *and* $SP_2(x_k, w)$ are the social polarization indices for factor $x_k$, and $GSP_\varepsilon = -\left(\frac{2}{n\bar{y}} \sum_{i=1}^{(n+1)/2} \varepsilon_i + \frac{GC_\varepsilon}{\bar{y}}\right)$.

Note: In this article we assume that *n* is odd. For *n* even, the decomposition is given by,

$$SP_1(y,w) = \sum_k \frac{\beta_{x_k} m_{x_k}}{m_y} SP_1(x_k, y) + \frac{2\bar{y}}{m_y} GSP_\varepsilon$$

$$SP_2(y,w) = \sum_k \frac{\beta_{x_k} \bar{x}_k}{\bar{y}} SP_2(x_k, w) + GSP_\varepsilon.$$

**Proof.** *See Appendix, Proof 3.*

These equations permit a partitioning of the causes of social polarization into an explained and an unexplained component, as in the decomposition of the concentration index in equation (1). In equations (12) and (13) the explained part is the sum of the contributions of each factor. The contribution of $x_k$ is equal to the elasticity times social polarization for $x_k$. One can notice that in (12) the elasticity is evaluated at the values taken by the individual with the median income, i.e. ($m_y$, $m_{x_k}$), whereas in the decomposition of the concentration index in equation (1) or in that of $SP_2$ in equation (13) it is evaluated at the mean ($\bar{y}$, $\bar{x}_k$). The presence of the term $SP(x_k, w)$ in the decompositions highlights that a factor $x_k$ has to be distributed in a polarized way across income levels for it to have an effect on social polarization in health.

The unexplained component of the decompositions is the residual part that cannot be explained by systematic variation in the $x_k$.

### 3.3. *Decomposition of Changes in Social Health Polarization*

In order to unravel the causes of the change of polarization between two subsamples $A$ and $B$, we can apply an Oaxaca-type decomposition (Oaxaca, 1973). If we denote by $\eta_{kA}$ and $\eta_{kB}$ the elasticities for the two subsamples, we get for both $SP_1$ and $SP_2$:

$$(14)\quad \Delta SP = \sum_k \eta_{kB}(SP_B(x_k, w) - SP_A(x_k, w)) + \sum_k SP_A(x_k, w)(\eta_{kB} - \eta_{kA}) + \Delta GSP.$$

An alternative decomposition is:

$$(15)\quad \Delta SP = \sum_k \eta_{kA}(SP_B(x_k, w) - SP_A(x_k, w)) + \sum_k SP_B(x_k, w)(\eta_{kB} - \eta_{kA}) + \Delta GSP.$$

## 4. Empirical Illustration

In the rest of the paper, we show that the social polarization approach complements the inequality measurement and we analyze the evolution of social polarization with age for French women.

### 4.1. *Data*

The data are taken from the French "Enquête Décennale Santé" which is a nationally representative health survey conducted in 2002–03. We restrict our sample to women because women's health distribution exhibits a polarization/depolarization phenomenon which is worth analyzing with the tools we have just presented, whereas men's distribution does not. After eliminating the observations with missing values we get a sample of 11,838 observations.

Since we are interested in the evolution of polarization with age, we create six age groups: 18 to 24, 25 to 34, 35 to 44, 45 to 54, 55 to 64, and more than 65. Every group contain between 1,257 and 2,465 women.

#### Factors

Before presenting the construction of the health variable, we concentrate on the factors that are correlated with health, because they are needed to create the health variable. These factors are the following:

- The income variable. We use an equivalized household income: total household income is divided by an equivalence factor in order to correct for the household size,[5] and expressed in logarithm in order to account for the concave relationship between health and income (Deaton and Paxson, 1998; Van Doorslaer and Jones, 2003). To keep in our sample women who

[5]Income by "unité de consommation."

declare no income, we add 1 to the equivalized income before taking the logarithm. So the individual income variable is $Ln(income_i + 1)$. Thus the vector of individual incomes $w = (w_1, \ldots, w_n)$ is equal in the empirical section to $(Ln(income_1 + 1), \ldots, Ln(income_n + 1))$.
- Three dummy variables indicating the education level: no education (reference); less than Baccalauréat; higher education.
- Dummy variables indicating the labor market status: employed (reference); unemployed; retired; not in the labor force (NLF).
- The marital status: married (reference); single; divorced or separated (Div/Sep); widowed.
- The number of children the woman gave birth to.
- The number of individuals in the household.
- The number of children in the household.
- Dummy variables for two events during childhood: handicap, accident or death of mother; money troubles.
- Dummy variables for two events last year: death of a friend or a member of the family ("Death last year"); money troubles.

### Health

To measure health, we use the predicted probability of reporting excellent or very good health. This probability is obtained after regressing a dichotomic subjective health variable on the factors presented above.

Subjective health comes from the survey question:

*"How is your health in general? Excellent, Very good, Good, Fair, Poor"*

We aggregate the "Excellent" and "Very good" categories on the one hand and the "Good," "Fair" and "Poor" categories on the other hand so as to obtain a dichotomic health variable indicating that individual health is excellent or very good.

Subjective health used in this paper has interesting properties. First, it offers a summary of an individual's general state of health. Second, many longitudinal studies have highlighted that a person's own appraisal of his or her health is a very good predictor of future mortality and morbidity. The correlation between subjective health and mortality remains strong, even after controlling for other health variables and for socioeconomic variables (see Idler and Benyamini, 1997, for a summary of this research).

An important challenge in measuring social inequality and polarization in health is to have a usable health measure. Indeed subjective health is ordinal and so it cannot be used directly to quantify inequality and polarization. In this context the traditional solution is to transform this ordinal variable into a cardinal one, by regressing the subjective health variable on a number of factors, then predictions of the linear index can be used as a measure of individual health after rescaling (Cutler *et al.*, 1997; Groot, 2000; Van Doorslaer and Jones, 2003).

In our analysis we focus on the probability of reporting excellent or very good health, using a linear probability model. We regress the dichotomic health variable (indicating excellent or very good health) on the factors,[6] and the predicted probability is our variable of interest. Our model takes the form:

$$h_i = \alpha + \sum_k \beta_k x_{ki} + \varepsilon_i$$

and the predicted probability $\hat{h}_i$ is the empiric analog of $y_i$.

The results of this model are given in Table 1. The first column reproduces the health specification described in the equation above for the whole sample, whereas the other columns show the results for the different age groups. The variable "retired" is excluded from the models for age groups 18–24 and 25–34 because there are no retired individuals under 34 years of age. For the same reason "Widow" for the age group 25–34 and "Unemployed" for the age group 65+ are excluded.

When we look at the first column, we observe that the probability of reporting excellent or very good health decreases with age as expected. On average more income is positively and significantly correlated with a larger probability of reporting excellent or very good health; similarly, money troubles during childhood or last year have a negative and significant effect on this probability. On the contrary education is positively and significantly related to this probability. Moreover unemployed and retired individuals and individuals who are not in the labor force have a smaller probability of reporting excellent or very good health than employed people. Interestingly we do not find any effect of the marital status on the probability. The number of births is negatively and significantly correlated with the probability of reporting excellent or very good health, whereas the number of children in the household is positively correlated with this probability.

### 4.2. *Comparison of Social Inequality and Polarization*

We now focus on the relationship between income and the probability of reporting excellent or very good health.

Figure 1 presents the bivariate densities for income and the probability for women. We observe that between ages 18–24 and 25–34 there is a clear polarization process. Then from age 25–34 to 45–54 the evolution is less clear. But the distribution polarizes with the formation of two peaks at age 55–64. Finally the densities suggest a depolarization phenomenon between ages 55–64 and 65+ since one of the two peaks diminishes.

We have calculated the concentration index $C$ and the social polarization index $SP_2$[7] for each age group and drawn their evolutions with age in Figure 2. Both indices are significantly positive for all age groups, which means that health is unequally distributed in favor of the higher income individuals. Moreover the inequality and polarization measures both support the idea of an increase of

[6]Our approach does not mean that the factors have a causal impact on health. Indeed, we do not control for possible endogeneity in the health–income relationship because our aim is simply to provide an empirical illustration of our methods. As a consequence some caution is required in giving a causality interpretation to our findings.

[7]For space reasons, we only report results for $SP_2$, but results for $SP_1$ are very similar.

TABLE 1
Regression Results by Age Groups

| | (1) All | (2) 18–24 | (3) 25–34 | (4) 35–44 | (5) 45–54 | (6) 55–64 | (7) 65+ |
|---|---|---|---|---|---|---|---|
| Ln(income+1) | 0.048 | 0.013 | 0.053 | 0.056 | 0.035 | 0.068 | 0.106 |
| | (6.90)*** | (1.49) | (3.30)*** | (3.31)*** | (1.92)* | (3.05)*** | (4.64)*** |
| Less than Baccalauréat | 0.074 | –0.016 | 0.008 | 0.062 | 0.098 | 0.141 | 0.041 |
| | (6.49)*** | (0.42) | (0.30) | (2.45)** | (3.49)*** | (4.47)*** | (1.64) |
| Higher education | 0.135 | 0.020 | 0.068 | 0.085 | 0.175 | 0.256 | 0.085 |
| | (10.75)*** | (0.56) | (2.48)** | (3.18)*** | (5.65)*** | (6.83)*** | (2.34)** |
| Unemployed | –0.054 | –0.049 | –0.043 | –0.035 | –0.125 | 0.003 | |
| | (3.35)*** | (1.68)* | (1.73)* | (1.05) | (3.46)*** | (0.05) | |
| Retired | –0.062 | | | –0.464 | –0.052 | –0.048 | –0.239 |
| | (3.71)*** | | | (2.71)*** | (0.68) | (1.72)* | (1.61) |
| NLF | –0.055 | 0.007 | –0.025 | –0.061 | –0.133 | –0.083 | –0.199 |
| | (5.05)*** | (0.33) | (1.20) | (2.79)*** | (5.36)*** | (2.60)*** | (1.31) |
| Single | –0.020 | 0.022 | –0.028 | –0.014 | 0.019 | –0.056 | –0.022 |
| | (1.36) | (0.86) | (1.02) | (0.36) | (0.38) | (0.85) | (0.39) |
| Widow | –0.026 | | –0.162 | –0.107 | 0.071 | –0.023 | –0.074 |
| | (1.41) | | (0.86) | (1.06) | (1.16) | (0.42) | (1.94)* |
| Div/sep | –0.013 | –0.256 | 0.041 | 0.021 | –0.009 | 0.016 | –0.048 |
| | (0.69) | (1.26) | (0.84) | (0.51) | (0.22) | (0.29) | (0.82) |
| No. births | –0.015 | –0.022 | 0.005 | –0.013 | –0.016 | –0.014 | –0.011 |
| | (4.25)*** | (0.98) | (0.34) | (0.98) | (1.74)* | (1.56) | (1.86)* |
| No. individuals in hh | –0.005 | 0.003 | 0.017 | 0.030 | 0.028 | 0.034 | –0.076 |
| | (0.45) | (0.15) | (0.73) | (0.96) | (0.86) | (0.89) | (2.28)** |
| No. children in hh | 0.033 | 0.012 | –0.007 | –0.001 | 0.005 | –0.028 | 0.040 |
| | (2.56)** | (0.57) | (0.24) | (0.04) | (0.16) | (0.59) | (0.73) |
| Hand., acci., death of mother dur. child. | –0.041 | –0.049 | 0.010 | –0.005 | –0.081 | –0.055 | –0.063 |
| | (2.91)*** | (1.57) | (0.32) | (0.16) | (2.27)** | (1.40) | (1.75)* |
| Money troubles dur. child. | –0.085 | –0.021 | –0.062 | –0.066 | –0.130 | –0.052 | –0.122 |
| | (7.08)*** | (0.68) | (2.53)** | (2.43)** | (4.80)*** | (1.56) | (3.79)*** |
| Death last year | –0.030 | –0.009 | 0.007 | –0.025 | –0.019 | –0.102 | –0.067 |
| | (2.80)*** | (0.38) | (0.37) | (1.18) | (0.77) | (3.06)*** | (2.01)** |
| Money troubles last year | –0.118 | –0.085 | –0.119 | –0.123 | –0.101 | –0.211 | –0.048 |
| | (7.90)*** | (2.67)*** | (4.66)*** | (4.36)*** | (3.06)*** | (4.06)*** | (0.73) |
| 25–34 | –0.056 | | | | | | |
| | (3.36)*** | | | | | | |
| 35–44 | –0.112 | | | | | | |
| | (6.47)*** | | | | | | |
| 45–54 | –0.160 | | | | | | |
| | (8.85)*** | | | | | | |
| 55–64 | –0.179 | | | | | | |
| | (8.70)*** | | | | | | |
| 65+ | –0.305 | | | | | | |
| | (12.60)*** | | | | | | |
| Constant | 0.411 | 0.755 | 0.307 | 0.166 | 0.298 | –0.101 | –0.117 |
| | (5.84)*** | (7.86)*** | (1.88)* | (0.93) | (1.50) | (0.42) | (0.42) |
| Observations | 11,838 | 1,257 | 2,051 | 2,465 | 2,317 | 1,629 | 2,119 |

*Notes*: Absolute value of t statistics in parentheses.
*significant at 10%; **significant at 5%; ***significant at 1%.
Reference categories: No education, Employed, Married, 18–24.
*Source*: Enquête Décennale Santé, with author's calculations.

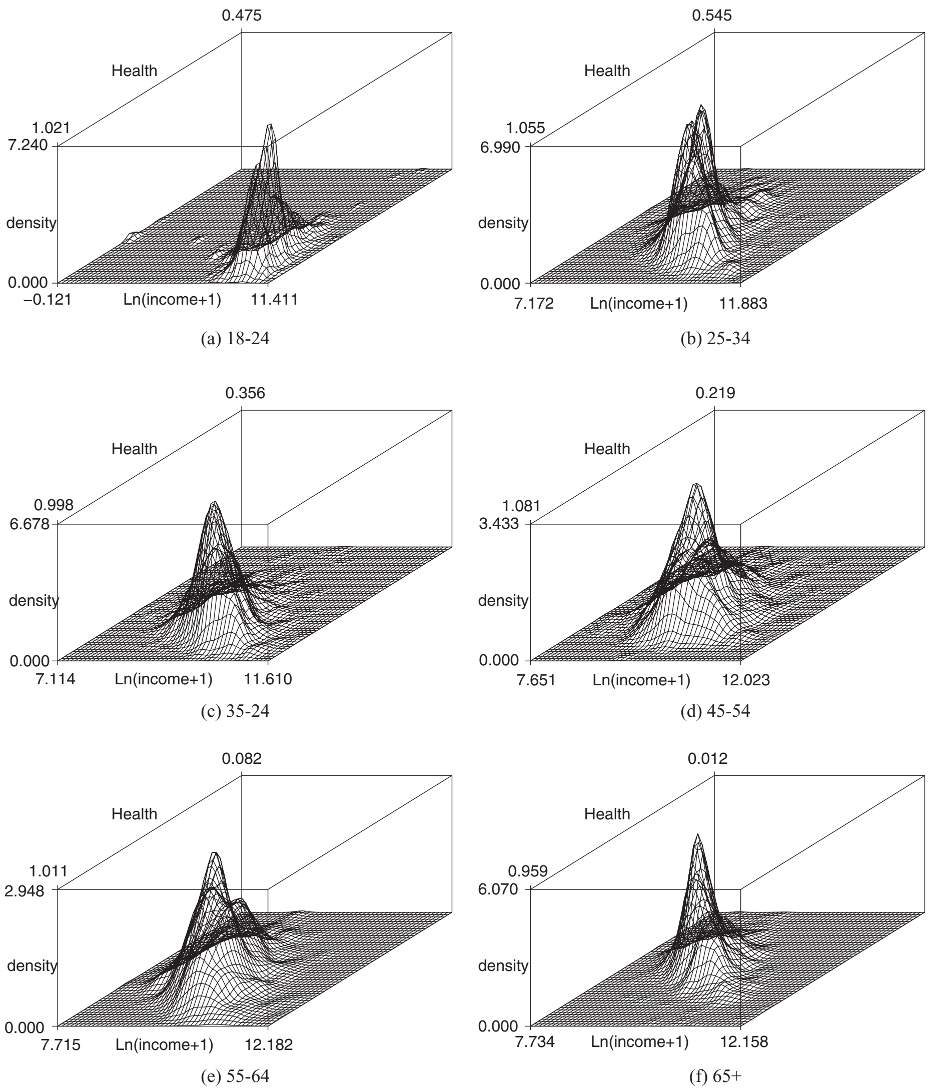


Figure 1. Density of Health and Income by Age Groups

*Source*: See Table 1.

dispersion between ages 18–24 and 25–34 and between 45–54 and 55–64. But interestingly inequality remains constant between ages 55–64 and 65+ whereas social polarization significantly decreases. As a conclusion our polarization measure $SP_2$ adequately shows the depolarization phenomenon that was previously observed, whereas this is not the case for the concentration index. So when densities suggest polarization or depolarization phenomena it seems that our social polarization measure is a useful complement to the concentration index.

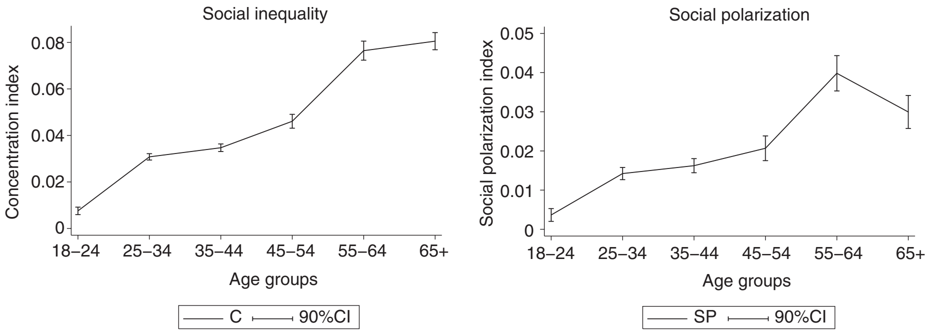


Figure 2. Social Inequality and Polarization

*Source*: See Table 1.

### 4.3. *Decomposition of Social Polarization Indices*

Describing the evolution with age of social polarization for the probability of reporting excellent or very good health is important, but more interesting for policy purposes is to quantify the contribution of various determinants of this probability to social polarization.

Since the contribution of a factor ($x_k$) to social polarization in the probability is the product of the factor elasticity of the probability $\left(\frac{\beta_{x_k}\bar{x}_k}{\bar{y}}\right)$ and the social polarization in the factor ($SP_2(x_k, w)$), requirements for a factor to have an impact on social polarization is that the factor elasticity of the probability is large and that the factor distribution across income levels is polarized. The decomposition of social polarization into the contributions of the factors is presented in Table 2. Columns (1) and (2) present the estimated elasticities and their standard errors; social polarization indices and their standard errors can be found in columns (3) and (4); the contributions of the factors and their standard errors are in columns (5) and (6); the last column contains the contribution of the factor expressed as a percentage of social polarization in the probability of reporting excellent or very good health. A factor's contribution can be either positive (when the factor is positively correlated with the probability) or negative (when the factor is negatively associated with the probability). In the last column, a p percent contribution of factor $x_k$ means that income-related health polarization would *ceteris paribus* be p percent lower if $x_k$ were equally distributed across income levels.

#### Income

The main result is that the contribution of income plays an important role in social polarization in health for all age groups. This is particularly true for women over 65 years old since for them social polarization in the probability of reporting excellent or very good health would be 85 percent lower if income was not correlated with the probability or if income was equally distributed. The contribution of income is decomposed into elasticity (column 1) and social polarization (column

TABLE 2

Decomposition Results by Age Groups

| Factor | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| | Elasticity | | Social Polarization of Factor | | Contribution of Factor | | Contribution of Factor in % |
| $x_k$ | $\frac{\beta_{x_k}\bar{x}_k}{\bar{y}}$ | SE | $SP_2(x_k, w)$ | SE | $\frac{\beta_{x_k}\bar{x}_k}{\bar{y}}SP_2(x_k, w)$ | SE | $100\frac{\beta_{x_k}\bar{x}_k}{\bar{y}}\frac{SP_2(x_k, w)}{SP_2(y, w)}$ |
| 18–24 | | | | | | | |
| Ln(income+1) | 0.1365 | 0.0004 | 0.0146 | 0.0004 | 0.002 | 0.0003 | 54.51 |
| Low education | –0.0047 | 0.0002 | –0.0463 | 0.0002 | 0.0002 | 0.0001 | 5.94 |
| Higher education | 0.015 | 0.0003 | 0.036 | 0.0003 | 0.0005 | 0.0002 | 14.80 |
| Unemployed | –0.0056 | 0.0005 | –0.2346 | 0.0005 | 0.0013 | 0.0003 | 35.97 |
| NLF | 0.0041 | 0.0001 | –0.0404 | 0.0001 | –0.0002 | 0.0001 | –4.53 |
| Single | 0.017 | 0.0003 | –0.034 | 0.0003 | –0.0006 | 0.0002 | –15.77 |
| Div/sep | –0.0004 | 0.0003 | –0.1266 | 0.0003 | 0.0001 | 0.0002 | 1.54 |
| No. births | –0.0032 | 0.0003 | –0.0803 | 0.0003 | 0.0003 | 0.0001 | 6.94 |
| No. individuals in hh | 0.009 | 0.0001 | –0.0108 | 0.0001 | –0.0001 | 0.0001 | –2.64 |
| No. children in hh | 0.0168 | 0.0005 | –0.0406 | 0.0005 | –0.0007 | 0.0003 | –18.66 |
| Hand., accident, death of mother dur. child. | –0.0039 | 0.0004 | –0.011 | 0.0004 | 0.0000 | 0.0002 | 1.18 |
| Money troubles dur. child. | –0.0019 | 0.0002 | 0.0465 | 0.0002 | –0.0001 | 0.0001 | –2.39 |
| Death last year | –0.0014 | 0.0001 | 0.0165 | 0.0001 | 0.0000 | 0.0001 | –0.65 |
| Money troubles last year | –0.0072 | 0.0007 | –0.1202 | 0.0007 | 0.0009 | 0.0004 | 23.74 |
| 25–34 | | | | | | | |
| Ln(income+1) | 0.5768 | 0.0008 | 0.0124 | 0.0008 | 0.0071 | 0.0003 | 50.08 |
| Low education | 0.0026 | 0.0001 | –0.1861 | 0.0001 | –0.0005 | 0.0000 | –3.46 |
| Higher education | 0.0489 | 0.0008 | 0.117 | 0.0008 | 0.0057 | 0.0004 | 40.18 |
| Unemployed | –0.0049 | 0.0003 | –0.1794 | 0.0003 | 0.0009 | 0.0002 | 6.20 |
| NLF | –0.0054 | 0.0002 | –0.2093 | 0.0002 | 0.0011 | 0.0001 | 7.91 |
| Single | –0.0062 | 0.0003 | –0.0225 | 0.0003 | 0.0001 | 0.0002 | 0.98 |
| Widow | –0.0003 | 0.0002 | –0.4602 | 0.0002 | 0.0001 | 0.0000 | 0.87 |
| Div/sep | 0.0013 | 0.0002 | –0.2211 | 0.0002 | –0.0003 | 0.0001 | –2.09 |
| No. births | 0.0065 | 0.0001 | –0.1042 | 0.0001 | –0.0007 | 0.0001 | –4.75 |

TABLE 2 (*continued*)

| Factor $x_k$ | (1) Elasticity $\frac{\beta_{x_k}\bar{x}_k}{\bar{y}}$ | (2) SE | (3) Social Polarization of Factor $SP_2(x_k, w)$ | (4) SE | (5) Contribution of Factor $\frac{\beta_{x_k}\bar{x}_k}{\bar{y}}SP_2(x_k, w)$ | (6) SE | (7) Contribution of Factor in % $100\frac{\beta_{x_k}\bar{x}_k}{\bar{y}}\frac{SP_2(x_k, w)}{SP_2(y, w)}$ |
|---|---|---|---|---|---|---|---|
| No. individuals in hh | 0.0589 | 0.0005 | –0.0387 | 0.0005 | –0.0023 | 0.0003 | –16.01 |
| No. children in hh | –0.0101 | 0.0002 | –0.1001 | 0.0002 | 0.001 | 0.0001 | 7.09 |
| Hand., accident, death of mother dur. child. | 0.0007 | 0.0001 | –0.0311 | 0.0001 | 0.0000 | 0.0000 | –0.14 |
| Money troubles dur. child. | –0.0071 | 0.0005 | –0.0756 | 0.0005 | 0.0005 | 0.0003 | 3.77 |
| Death last year | 0.0013 | 0.0001 | –0.0356 | 0.0001 | 0.0000 | 0.0000 | –0.33 |
| Money troubles last year | –0.0132 | 0.0009 | –0.1045 | 0.0009 | 0.0014 | 0.0005 | 9.67 |
| 35–44 | | | | | | | |
| Ln(income+1) | 0.6594 | 0.0011 | 0.0129 | 0.0011 | 0.0085 | 0.0004 | 52.26 |
| Low education | 0.0301 | 0.0008 | –0.0763 | 0.0008 | –0.0023 | 0.0004 | –14.13 |
| Higher education | 0.05 | 0.001 | 0.1101 | 0.001 | 0.0055 | 0.0005 | 33.88 |
| Unemployed | –0.0026 | 0.0002 | –0.1652 | 0.0002 | 0.0004 | 0.0001 | 2.64 |
| Retired | –0.0012 | 0.0005 | –0.3995 | 0.0005 | 0.0005 | 0.0002 | 2.83 |
| NLF | –0.0128 | 0.0006 | –0.1402 | 0.0006 | 0.0018 | 0.0003 | 11.03 |
| Single | –0.0018 | 0.0001 | 0.0001 | 0.0001 | 0.0000 | 0.0001 | –0.00 |
| Widow | –0.0008 | 0.0002 | –0.0109 | 0.0002 | 0.0000 | 0.0001 | 0.05 |
| Div/sep | 0.0021 | 0.0001 | –0.1327 | 0.0001 | –0.0003 | 0.0001 | –1.72 |
| No. births | –0.0308 | 0.0004 | –0.0347 | 0.0004 | 0.0011 | 0.0002 | 6.57 |
| No. individuals in hh | 0.1347 | 0.0009 | –0.0151 | 0.0009 | –0.002 | 0.0006 | –12.51 |
| No. children in hh | –0.0032 | 0.0000 | –0.0343 | 0.0000 | 0.0001 | 0.0000 | 0.68 |
| Hand., accident, death of mother dur. child. | –0.0004 | 0.0000 | –0.0006 | 0.0000 | 0.0000 | 0.0000 | 0.00 |
| Money troubles dur. child. | –0.0075 | 0.0005 | –0.0413 | 0.0005 | 0.0003 | 0.0003 | 1.90 |
| Death last year | –0.0047 | 0.0002 | –0.016 | 0.0002 | 0.0001 | 0.0001 | 0.46 |
| Money troubles last year | –0.014 | 0.0009 | –0.1862 | 0.0009 | 0.0026 | 0.0005 | 16.02 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 45–54 | | | | | | | |
| Ln(income+1) | 0.4626 | 0.0014 | 0.0134 | 0.0014 | 0.0062 | 0.0003 | 29.87 |
| Low education | 0.0666 | 0.0014 | –0.0371 | 0.0014 | –0.0025 | 0.0008 | –11.95 |
| Higher education | 0.0871 | 0.0022 | 0.1121 | 0.0022 | 0.0098 | 0.0012 | 47.22 |
| Unemployed | –0.0116 | 0.0009 | –0.1656 | 0.0009 | 0.0019 | 0.0005 | 9.32 |
| Retired | –0.001 | 0.0002 | 0.2462 | 0.0002 | –0.0002 | 0.0001 | –1.15 |
| NLF | –0.0309 | 0.0015 | –0.1491 | 0.0015 | 0.0046 | 0.0008 | 22.30 |
| Single | 0.0014 | 0.0001 | –0.0007 | 0.0001 | 0.0000 | 0.0001 | –0.00 |
| Widow | 0.0026 | 0.0003 | –0.0464 | 0.0003 | –0.0001 | 0.0002 | –0.59 |
| Div/sep | –0.0015 | 0.0001 | –0.0906 | 0.0001 | 0.0001 | 0.0000 | 0.65 |
| No. births | –0.0467 | 0.0006 | –0.0313 | 0.0006 | 0.0015 | 0.0003 | 7.07 |
| No. individuals in hh | 0.1059 | 0.0009 | –0.0203 | 0.0009 | –0.0021 | 0.0005 | –10.37 |
| No. children in hh | 0.0069 | 0.0002 | –0.0652 | 0.0002 | –0.0004 | 0.0001 | –2.16 |
| Hand., accident, death of mother dur. child. | –0.0073 | 0.0006 | –0.0412 | 0.0006 | 0.0003 | 0.0003 | 1.44 |
| Money troubles dur. child. | –0.0224 | 0.0013 | –0.0219 | 0.0013 | 0.0005 | 0.0007 | 2.36 |
| Death last year | –0.004 | 0.0002 | 0.0135 | 0.0002 | –0.0001 | 0.0001 | –0.25 |
| Money troubles last year | –0.0121 | 0.0008 | –0.1067 | 0.0008 | 0.0013 | 0.0004 | 6.21 |
| 55–64 | | | | | | | |
| Ln(income+1) | 0.9974 | 0.0046 | 0.0136 | 0.0046 | 0.0135 | 0.0007 | 33.98 |
| Low education | 0.1113 | 0.0027 | –0.0261 | 0.0027 | –0.0029 | 0.0015 | –7.29 |
| Higher education | 0.1119 | 0.004 | 0.1791 | 0.0040 | 0.0200 | 0.0022 | 50.34 |
| Unemployed | 0.0003 | 0.0000 | –0.1284 | 0.0000 | 0.0000 | 0.0000 | –0.08 |
| Retired | –0.0274 | 0.0009 | 0.029 | 0.0009 | –0.0008 | 0.0005 | –1.99 |
| NLF | –0.0293 | 0.0014 | –0.0843 | 0.0014 | 0.0025 | 0.0008 | 6.20 |
| Single | –0.0044 | 0.0005 | 0.0622 | 0.0005 | –0.0003 | 0.0003 | –0.68 |
| Widow | –0.0031 | 0.0002 | –0.1249 | 0.0002 | 0.0004 | 0.0001 | 0.96 |
| Div/sep | 0.0024 | 0.0002 | –0.0455 | 0.0002 | –0.0001 | 0.0001 | –0.26 |
| No. births | –0.0471 | 0.0008 | –0.051 | 0.0008 | 0.0024 | 0.0004 | 6.03 |
| No. individuals in hh | 0.1054 | 0.0011 | –0.0078 | 0.0011 | –0.0008 | 0.0006 | –2.05 |
| No. children in hh | –0.0096 | 0.0006 | –0.0728 | 0.0006 | 0.0007 | 0.0003 | 1.75 |
| Hand., accident, death of mother dur. child. | –0.0076 | 0.0006 | –0.081 | 0.0006 | 0.0006 | 0.0003 | 1.55 |

TABLE 2 (*continued*)

| Factor $x_k$ | (1) Elasticity $\frac{\beta_{x_k}\bar{x}_k}{\bar{y}}$ | (2) SE | (3) Social Polarization of Factor $SP_2(x_k, w)$ | (4) SE | (5) Contribution of Factor $\frac{\beta_{x_k}\bar{x}_k}{\bar{y}}SP_2(x_k, w)$ | (6) SE | (7) Contribution of Factor in % $100\frac{\beta_{x_k}\bar{x}_k}{\bar{y}}\frac{SP_2(x_k, w)}{SP_2(y, w)}$ |
|---|---|---|---|---|---|---|---|
| Money troubles dur. child. | –0.0108 | 0.0007 | –0.0389 | 0.0007 | 0.0004 | 0.0004 | 1.05 |
| Death last year | –0.0206 | 0.0013 | 0.0233 | 0.0013 | –0.0005 | 0.0008 | –1.20 |
| Money troubles last year | –0.017 | 0.0018 | –0.2745 | 0.0018 | 0.0047 | 0.0009 | 11.68 |
| 65+ | | | | | | | |
| Ln(income+1) | 2.0957 | 0.0087 | 0.0121 | 0.0087 | 0.0254 | 0.0012 | 84.87 |
| Low education | 0.045 | 0.0009 | 0.0143 | 0.0009 | 0.0006 | 0.0005 | 2.14 |
| Higher education | 0.0301 | 0.0014 | 0.1806 | 0.0014 | 0.0054 | 0.0007 | 18.11 |
| Retired | –0.4137 | 0.0045 | 0.0139 | 0.0045 | –0.0058 | 0.0024 | –19.19 |
| NLF | –0.0664 | 0.0033 | –0.0716 | 0.0033 | 0.0048 | 0.002 | 15.87 |
| Single | –0.0026 | 0.0002 | 0.0265 | 0.0002 | –0.0001 | 0.0001 | –0.23 |
| Widow | –0.0592 | 0.0017 | –0.0255 | 0.0017 | 0.0015 | 0.0009 | 5.03 |
| Div/sep | –0.0054 | 0.0005 | 0.0349 | 0.0005 | –0.0002 | 0.0003 | –0.62 |
| No. births | –0.0602 | 0.0011 | –0.0247 | 0.0011 | 0.0015 | 0.0005 | 4.95 |
| No. individuals in hh | –0.2607 | 0.003 | 0.0087 | 0.003 | –0.0023 | 0.0014 | –7.53 |
| No. children in hh | 0.0064 | 0.0006 | 0.0291 | 0.0006 | 0.0002 | 0.0003 | 0.62 |
| Hand., accident, death of mother dur. child. | –0.0127 | 0.0009 | 0.0343 | 0.0009 | –0.0004 | 0.0005 | –1.45 |
| Money troubles dur. child. | –0.0338 | 0.0019 | 0.0026 | 0.0019 | –0.0001 | 0.001 | –0.29 |
| Death last year | –0.0163 | 0.001 | 0.051 | 0.001 | –0.0008 | 0.0006 | –2.76 |
| Money troubles last year | –0.0029 | 0.0004 | –0.0508 | 0.0004 | 0.0001 | 0.0002 | 0.48 |

*Note*: Standard errors were computed using Stata nlcom command.
*Source*: See Table 1.

3). In order to give an accurate interpretation of the results, one has to notice that "Ln(income+1)" is the variable $w$ which gives us the ranking of the individuals. So the index of social polarization of income $SP_2(w, w)$ can be expressed using Wolfson's polarization index, $SP_2(w, w) = \frac{m(w)}{2\overline{w}} P(w)$.[8] Consequently $SP_2(w, w)$ is a measure of univariate polarization in income. Income polarization is positive and significant for all age groups (except for the age group 65+), whereas the income elasticity of health is positive, large, and significant. So except for women above 65, both the elasticity and the polarization of income explain the size of the contribution of income to social polarization of health.

#### Education

The contribution of higher education is also important, since this explains between 14.80 percent and 50.34 percent of social polarization in the probability of reporting excellent or very good health. For all age groups both the elasticity and the income-related polarization in higher education are positive and significant. This corresponds to the positive correlation between higher education and health found in Table 1, but it also suggests that the distribution of education across income levels is polarized.

#### Number of Individuals in the Household

Between 25 and 55 years of age, the family size has a negative and significant contribution to social polarization in the probability for women. The elasticity is positive and significant, whereas social polarization in this factor is negative.

### 4.4. *Decomposition of the Change of Social Polarization*

We now turn to the more important question: how do the factors contribute to the changes in polarization with age? The Oaxaca-type decompositions permit us to answer this question. The empiric analogs of equations (14) and (15) are provided in Table 3.

#### The Polarization Process Between Ages 18–24 and 25–34

As shown in Figure 2, $SP_2$ significantly increases between ages 18–24 and 25–34. Table 3 suggests that changes in the contributions of income and higher education are the two main explanations, and that the two impacts are very similar in magnitude. The increase of the contribution of income is above all due to the increase of the income elasticity of health, and not to a change of the income polarization, whereas the strengthening of the contribution of higher education reflects an increase of both the elasticity and the social polarization of higher education. The number of children in the household also plays an important part in the rise of polarization between ages 18–24 and 25–34.

[8]Similarly $SP_1(w, w) = P(w)$.

TABLE 3
Oaxaca-type Decomposition of Changes of Social Polarization in Health

| Factor | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
|---|---|---|---|---|---|---|---|---|---|---|
| | Equation (14) | | | | Equation (15) | | | | Total | |
| $x_k$ | $\eta\Delta SP_2$ | SE | $SP_2\Delta\eta$ | SE | $\eta\Delta SP_2$ | SE | $SP_2\Delta\eta$ | SE | $\Delta(\eta SP_2)$ | SE |
| Between 18–24 and 25–34 | | | | | | | | | | |
| Ln(income+1) | –0.0013 | 0.0013 | 0.0064 | 0.0009 | –0.0003 | 0.0003 | 0.0054 | 0.0003 | 0.0051 | 0.0005 |
| Low education | –0.0004 | 0.0001 | –0.0003 | 0.0002 | 0.0007 | 0.0002 | –0.0014 | 0.0001 | –0.0007 | 0.0001 |
| Higher education | 0.0040 | 0.0007 | 0.0012 | 0.0004 | 0.0012 | 0.0002 | 0.0040 | 0.0003 | 0.0052 | 0.0005 |
| Unemployed | –0.0003 | 0.0003 | –0.0002 | 0.0001 | –0.0003 | 0.0003 | –0.0001 | 0.0001 | –0.0004 | 0.0003 |
| NLF | 0.0009 | 0.0001 | 0.0004 | 0.0001 | –0.0007 | 0.0001 | 0.0020 | 0.0002 | 0.0013 | 0.0001 |
| Single | –0.0001 | 0.0002 | 0.0008 | 0.0002 | 0.0002 | 0.0005 | 0.0005 | 0.0006 | 0.0007 | 0.0002 |
| Div/sep | –0.0001 | 0.0006 | –0.0002 | 0.0008 | 0.0000 | 0.0002 | –0.0004 | 0.0001 | –0.0004 | 0.0002 |
| No. births | –0.0002 | 0.0003 | –0.0008 | 0.0004 | 0.0001 | 0.0002 | –0.001 | 0.0001 | –0.0009 | 0.0002 |
| No. individuals in hh | –0.0016 | 0.0006 | –0.0005 | 0.0004 | –0.0002 | 0.0001 | –0.0019 | 0.0003 | –0.0022 | 0.0003 |
| No. children in hh | 0.0006 | 0.0002 | 0.0011 | 0.0005 | –0.001 | 0.0004 | 0.0027 | 0.0003 | 0.0017 | 0.0003 |
| Hand., accident, death of mother dur. child. | 0.0000 | 0.0001 | –0.0001 | 0.0003 | 0.0001 | 0.0003 | –0.0001 | 0.0002 | –0.0001 | 0.0002 |
| Money troubles dur. child. | 0.0009 | 0.0004 | –0.0002 | 0.0003 | 0.0002 | 0.0001 | 0.0004 | 0.0002 | 0.0006 | 0.0003 |
| Death last year | –0.0001 | 0.0001 | 0.0000 | 0.0001 | 0.0001 | 0.0001 | –0.0001 | 0.0001 | 0.0000 | 0.0001 |
| Money troubles last year | –0.0002 | 0.0009 | 0.0007 | 0.0004 | –0.0001 | 0.0005 | 0.0006 | 0.0002 | 0.0005 | 0.0006 |
| Between 25–34 and 35–44 | | | | | | | | | | |
| Ln(income+1) | 0.0003 | 0.0005 | 0.0010 | 0.0001 | 0.0003 | 0.0005 | 0.0011 | 0.0000 | 0.0014 | 0.0005 |
| Low education | 0.0033 | 0.0007 | –0.0051 | 0.0005 | 0.0003 | 0.0001 | –0.0021 | 0.0004 | –0.0018 | 0.0004 |
| Higher education | –0.0003 | 0.0007 | 0.0001 | 0.0001 | –0.0003 | 0.0006 | 0.0001 | 0.0001 | –0.0002 | 0.0007 |
| Unemployed | 0.0000 | 0.0002 | –0.0004 | 0.0001 | –0.0001 | 0.0003 | –0.0004 | 0.0001 | –0.0005 | 0.0002 |
| NLF | –0.0009 | 0.0005 | 0.0015 | 0.0003 | –0.0004 | 0.0002 | 0.001 | 0.0002 | 0.0007 | 0.0003 |
| Single | 0.0000 | 0.0001 | –0.0001 | 0.0001 | –0.0001 | 0.0003 | 0.0000 | 0.0002 | –0.0001 | 0.0002 |
| Widow | –0.0004 | 0.0003 | 0.0003 | 0.0003 | –0.0001 | 0.0000 | 0.0000 | 0.0001 | –0.0001 | 0.0001 |
| Div/sep | 0.0002 | 0.0002 | –0.0002 | 0.0001 | 0.0001 | 0.0001 | –0.0001 | 0.0000 | 0.0000 | 0.0001 |
| No. births | –0.0021 | 0.0005 | 0.0039 | 0.0005 | 0.0005 | 0.0001 | 0.0013 | 0.0003 | 0.0017 | 0.0002 |
| No. individuals in hh | 0.0032 | 0.0009 | –0.0029 | 0.0004 | 0.0014 | 0.0004 | –0.0011 | 0.0003 | 0.0002 | 0.0006 |
| No. children in hh | –0.0002 | 0.0000 | –0.0007 | 0.0001 | –0.0007 | 0.0001 | –0.0002 | 0.0001 | –0.0009 | 0.0001 |
| Hand., accident, death of mother dur. child. | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Money troubles dur. child. | –0.0003 | 0.0004 | 0.0000 | 0.0001 | –0.0002 | 0.0004 | 0.0000 | 0.0000 | –0.0002 | 0.0004 |
| Death last year | –0.0001 | 0.0002 | 0.0002 | 0.0002 | 0.0000 | 0.0001 | 0.0001 | 0.0002 | 0.0001 | 0.0001 |
| Money troubles last year | 0.0011 | 0.0007 | 0.0001 | 0.0001 | 0.0011 | 0.0007 | 0.0001 | 0.0002 | 0.0012 | 0.0007 |
| Between 35–44 and 45–54 | | | | | | | | | | |
| Ln(income+1) | 0.0002 | 0.0004 | –0.0025 | 0.0001 | 0.0003 | 0.0005 | –0.0026 | 0.0001 | –0.0023 | 0.0004 |
| Low education | 0.0026 | 0.0012 | –0.0028 | 0.0005 | 0.0012 | 0.0005 | –0.0014 | 0.0004 | –0.0002 | 0.0009 |
| Higher education | 0.0002 | 0.0016 | 0.0041 | 0.0005 | 0.0001 | 0.0009 | 0.0042 | 0.0006 | 0.0043 | 0.0014 |
| Unemployed | 0.0000 | 0.0008 | 0.0015 | 0.0005 | 0.0000 | 0.0002 | 0.0015 | 0.0004 | 0.0015 | 0.0005 |
| Retired | –0.0006 | 0.0002 | –0.0001 | 0.0002 | –0.0007 | 0.0003 | 0.0000 | 0.0001 | –0.0007 | 0.0002 |
| NLF | 0.0003 | 0.0012 | 0.0025 | 0.0006 | 0.0001 | 0.0005 | 0.0027 | 0.0005 | 0.0028 | 0.0009 |
| Single | 0.0000 | 0.0001 | 0.0000 | 0.0001 | 0.0000 | 0.0001 | 0.0000 | 0.0002 | 0.0000 | 0.0001 |
| Widow | –0.0001 | 0.0004 | 0.0000 | 0.0005 | 0.0000 | 0.0001 | –0.0002 | 0.0002 | –0.0001 | 0.0002 |
| Div/sep | –0.0001 | 0.0001 | 0.0005 | 0.0001 | 0.0001 | 0.0001 | 0.0003 | 0.0001 | 0.0004 | 0.0001 |
| No. births | –0.0002 | 0.0005 | 0.0006 | 0.0001 | –0.0001 | 0.0003 | 0.0005 | 0.0001 | 0.0004 | 0.0004 |
| No. individuals in hh | –0.0005 | 0.0007 | 0.0004 | 0.0001 | –0.0007 | 0.0009 | 0.0006 | 0.0002 | –0.0001 | 0.0008 |
| No. children in hh | –0.0002 | 0.0001 | –0.0003 | 0.0001 | 0.0001 | 0.0000 | –0.0007 | 0.0001 | –0.0006 | 0.0001 |
| Hand., accident, death of mother dur. child. | 0.0003 | 0.0005 | 0.0000 | 0.0003 | 0.0000 | 0.0000 | 0.0003 | 0.0003 | 0.0003 | 0.0003 |
| Money troubles dur. child. | –0.0004 | 0.0011 | 0.0006 | 0.0005 | –0.0001 | 0.0004 | 0.0003 | 0.0005 | 0.0002 | 0.0008 |
| Death last year | –0.0001 | 0.0002 | 0.0000 | 0.0000 | –0.0001 | 0.0002 | 0.0000 | 0.0000 | –0.0001 | 0.0002 |
| Money troubles last year | –0.001 | 0.0007 | –0.0004 | 0.0002 | –0.0011 | 0.0007 | –0.0002 | 0.0002 | –0.0013 | 0.0006 |
| Between 45–54 and 55–64 | | | | | | | | | | |
| Ln(income+1) | 0.0002 | 0.0009 | 0.0071 | 0.0003 | 0.0001 | 0.0004 | 0.0073 | 0.0004 | 0.0074 | 0.0007 |
| Low education | 0.0012 | 0.002 | –0.0017 | 0.0005 | 0.0007 | 0.0012 | –0.0012 | 0.0006 | –0.0004 | 0.0017 |
| Higher education | 0.0075 | 0.0028 | 0.0028 | 0.0007 | 0.0058 | 0.0022 | 0.0044 | 0.0008 | 0.0103 | 0.0025 |
| Unemployed | 0.0000 | 0.0000 | –0.002 | 0.0005 | –0.0004 | 0.0008 | –0.0015 | 0.0006 | –0.002 | 0.0005 |
| Retired | 0.0059 | 0.0029 | –0.0065 | 0.0028 | 0.0002 | 0.0001 | –0.0008 | 0.0005 | –0.0006 | 0.0005 |
| NLF | –0.0019 | 0.0011 | –0.0002 | 0.0003 | –0.002 | 0.0012 | –0.0001 | 0.0002 | –0.0021 | 0.0011 |
| Single | –0.0003 | 0.0003 | 0.0000 | 0.0003 | 0.0001 | 0.0001 | –0.0004 | 0.0004 | –0.0003 | 0.0003 |
| Widow | 0.0002 | 0.0003 | 0.0003 | 0.0004 | –0.0002 | 0.0002 | 0.0007 | 0.0002 | 0.0005 | 0.0002 |
| Div/sep | 0.0001 | 0.0001 | –0.0003 | 0.0001 | –0.0001 | 0.0001 | –0.0002 | 0.0002 | –0.0002 | 0.0001 |
| No. births | 0.0009 | 0.0005 | 0.0000 | 0.0000 | 0.0009 | 0.0005 | 0.0000 | 0.0001 | 0.0009 | 0.0005 |
| No. individuals in hh | 0.0013 | 0.0008 | 0.0000 | 0.0000 | 0.0013 | 0.0008 | 0.0000 | 0.0000 | 0.0013 | 0.0008 |
| No. children in hh | 0.0001 | 0.0004 | 0.0011 | 0.0002 | –0.0001 | 0.0003 | 0.0012 | 0.0006 | 0.0011 | 0.0004 |
| Hand., accident, death of mother dur. child. | 0.0003 | 0.0005 | 0.0000 | 0.0000 | 0.0003 | 0.0005 | 0.0000 | 0.0001 | 0.0003 | 0.0005 |
| Money troubles dur. child. | 0.0002 | 0.0005 | –0.0003 | 0.0004 | 0.0004 | 0.0011 | –0.0004 | 0.0004 | –0.0001 | 0.0008 |
| Death last year | –0.0002 | 0.001 | –0.0002 | 0.0005 | 0.0000 | 0.0002 | –0.0004 | 0.0006 | –0.0004 | 0.0008 |
| Money troubles last year | 0.0028 | 0.0011 | 0.0005 | 0.0003 | 0.002 | 0.0009 | 0.0013 | 0.0005 | 0.0034 | 0.001 |

TABLE3 (*continued*)

| Factor | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
|---|---|---|---|---|---|---|---|---|---|---|
| | Equation (14) | | | | Equation (15) | | | | Total | |
| $x_k$ | $\eta\Delta SP_2$ | SE | $SP_2\Delta\eta$ | SE | $\eta\Delta SP_2$ | SE | $SP_2\Delta\eta$ | SE | $\Delta(\eta SP_2)$ | SE |
| Between 55–64 and 65+ | | | | | | | | | | |
| Ln(income+1) | –0.003 | 0.0018 | 0.0149 | 0.0008 | –0.0014 | 0.0009 | 0.0133 | 0.0006 | 0.0119 | 0.0014 |
| Low education | 0.0018 | 0.0008 | 0.0017 | 0.0009 | 0.0045 | 0.002 | –0.001 | 0.0008 | 0.0035 | 0.0016 |
| Higher education | 0.0000 | 0.0011 | –0.0147 | 0.0016 | 0.0002 | 0.004 | –0.0148 | 0.0026 | –0.0146 | 0.0023 |
| Retired | 0.0062 | 0.0078 | –0.0112 | 0.0069 | 0.0004 | 0.0005 | –0.0054 | 0.0023 | –0.005 | 0.0025 |
| NLF | –0.0008 | 0.0026 | 0.0031 | 0.001 | –0.0004 | 0.0012 | 0.0027 | 0.0011 | 0.0023 | 0.0021 |
| Single | 0.0001 | 0.0002 | 0.0001 | 0.0001 | 0.0002 | 0.0003 | 0.0000 | 0.0001 | 0.0002 | 0.0003 |
| Widow | –0.0059 | 0.0027 | 0.007 | 0.0024 | –0.0003 | 0.0001 | 0.0014 | 0.0009 | 0.0011 | 0.0009 |
| Div/sep | –0.0004 | 0.0004 | 0.0004 | 0.0003 | 0.0002 | 0.0002 | –0.0003 | 0.0004 | –0.0001 | 0.0003 |
| No. births | –0.0016 | 0.0008 | 0.0007 | 0.0002 | –0.0012 | 0.0006 | 0.0003 | 0.0001 | –0.0009 | 0.0007 |
| No. individuals in hh | –0.0043 | 0.002 | 0.0028 | 0.0021 | 0.0017 | 0.0008 | –0.0032 | 0.0019 | –0.0015 | 0.0015 |
| No. children in hh | 0.0007 | 0.0004 | –0.0012 | 0.0006 | –0.001 | 0.0006 | 0.0005 | 0.0008 | –0.0005 | 0.0005 |
| Hand., accident, death of mother dur. child. | –0.0015 | 0.0008 | 0.0004 | 0.0003 | –0.0009 | 0.0004 | –0.0002 | 0.0002 | –0.0011 | 0.0006 |
| Money troubles dur. child. | –0.0014 | 0.0016 | 0.0009 | 0.0008 | –0.0004 | 0.0005 | –0.0001 | 0.0007 | –0.0005 | 0.0011 |
| Death last year | –0.0004 | 0.0008 | 0.0001 | 0.0002 | –0.0006 | 0.0011 | 0.0002 | 0.0002 | –0.0003 | 0.001 |
| Money troubles last year | –0.0006 | 0.0003 | –0.0039 | 0.0008 | –0.0038 | 0.0016 | –0.0007 | 0.0011 | –0.0045 | 0.0009 |

*Note*: Standard errors were computed using Stata nlcom command.
*Source*: See Table 1.

### The Polarization Phenomenon Between Ages 45–54 and 55–64

The second polarization episode, which takes place between ages 45–54 and 55–64, can also be explained by the increase of the contribution of income and higher education. Again the increase of the income elasticity of health and of the social polarization in higher education are the main explanations.

Other factors also play a significant role: being a widow, the number of births, the number of children in the household, and money troubles the year before. However their contributions are small.

### The Depolarization Episode Between Ages 55–64 and 65+

Figure 2 highlights that social polarization diminishes between ages 55–64 and 65+. This may be due to a generation effect and to early mortality of the poorest women, which are not taken into account in our model.

However Table 3 sheds some light on this evolution and suggests that the depolarization process between ages 55–64 and 65+ is due to changes in respect of higher education and retirement (columns (9) and (10)). For higher education, the change in elasticity is the main explanation for the change in the contribution to $SP_2$. This finding has to be related to the decreases of the effects of higher education (compared to no education) on the probability of reporting excellent or very good health between ages 55–64 and 65+ in Table 1.

The changes of the contributions of the factors "handicap, accident, death of mother during childhood" and "money troubles last year" are also making for less social polarization in the probability of reporting excellent or very good health for women over 65.

On the contrary changes in respect of income are making for more social polarization in health. This is due to the increase of the correlation between income and the probability of reporting excellent or very good health. So these changes tend to diminish the depolarization process.

Changes in respect of a number of variables are negligible, in particular the marital status, the number of births, of children and individuals in the household, and of money troubles during childhood.

## 5. Conclusion

Our main aim in this paper has been to develop two original measures to quantify bivariate polarization. Bivariate polarization in a variable of interest is caused by bivariate polarization in the determinants of the variable of interest, and a decomposition method allows one to assess the importance of these different determinants in generating bivariate polarization in the variable of interest. Moreover changes in bivariate polarization can also be decomposed into changes in the contributions of the various determinants.

We then examined polarization in the probability of reporting excellent or very good health across income levels using French data on women. We found that income-related polarization in the probability of reporting excellent or very good health increases or remains stable for women between 18 and 64 years of age and decreases after age 65. On the contrary social inequality remains constant for women over 65. This makes clear that polarization measures convey additional information to that contained in social inequality. As a consequence the use of bivariate polarization indices in health economics is empirically relevant. Thanks to the decomposition techniques, the empirical section also explained the reasons for the changes of social polarization in health with age. The main findings are that the polarization episodes between ages 18–24 and 25–34 and between 45–54 and 55–64 are mainly due to increases in the contributions of income and higher education, whereas the depolarization that occurs after 65 reflects a large reduction in the contribution of higher education. Finally even if our results were intended to illustrate the use of the new measures, they shed some light on the reasons of the rise and decline of polarization with age for French women.

## Appendix

**Proof 1.**

Proof that $SP_1$ and $SP_2$ satisfy the increased spread axiom:

By equation (9),

$$SP_2(y^1, w) = \frac{2}{n\bar{y}^1}\sum_{k=1}^{n}(2y_k^{1M} - y_k^1)R_{w,k} - 1.$$

Assume that $(y^2, w)$ is derived from $(y^1, w)$ via a strict equalizing transfer between individuals $i$ and $j$ such that $i \le \frac{n-1}{2}$ and $j \ge \frac{n+3}{2}$. Then:

- $\bar{y}^1 = \bar{y}^2$
- For $k \le \frac{n-1}{2}$ and $k \ne i$,
  $y_k^{1M} = y_k^{2M} - \frac{2}{n-1}(y_j^1 - y_i^1)$ and $y_k^1 = y_k^2$
  So $2y_k^{1M} - y_k^1 = (2y_k^{2M} - y_k^2) - \frac{4}{n-1}(y_j^1 - y_i^1)$
- $y_i^{1M} = y_i^{2M} - \frac{2}{n-1}(y_j^1 - y_i^1)$, $y_i^1 = y_j^2$ and $y_j^1 = y_i^2$
  So $2y_i^{1M} - y_i^1 = (2y_i^{2M} - y_i^2) - \frac{4}{n-1}(y_j^1 - y_i^1) + (y_j^1 - y_i^1) = (2y_i^{2M} - y_i^2) + \frac{n-5}{n-1}(y_j^1 - y_i^1)$.
- For $k \ge \frac{n+1}{2}$ and $k \ne j$,
  $y_k^{1M} = y_k^{2M} + \frac{2}{n+1}(y_j^1 - y_i^1)$ and $y_k^1 = y_k^2$
  So $2y_k^{1M} - y_k^1 = (2y_k^{2M} - y_k^2) + \frac{4}{n+1}(y_j^1 - y_i^1)$.
- $y_j^{1M} = y_j^{2M} + \frac{2}{n+1}(y_j^1 - y_i^1)$, $y_i^1 = y_j^2$ and $y_j^1 = y_i^2$
  So $2y_j^{1M} - y_j^1 = (2y_j^{2M} - y_j^2) + \frac{4}{n+1}(y_j^1 - y_i^1) - (y_j^1 - y_i^1) = (2y_j^{2M} - y_j^2) - \frac{n-3}{n+1}(y_j^1 - y_i^1)$.

Consequently,

$$SP_2(y^1,w)=SP_2(y^2,w)$$
$$+\frac{2}{n\bar{y}^1}\left(y_j^1-y_i^1\right)\left[-\frac{4}{n-1}\sum_{k=1,k\neq i}^{\frac{n-1}{2}}R_{w,k}+\frac{n-5}{n-1}R_{w,i}+\frac{4}{n+1}\sum_{k=\frac{n+1}{2},k\neq j}^{n}R_{w,k}-\frac{n-3}{n+1}R_{w,j}\right]$$

Now $R_{w,k}=\frac{2k-1}{2n}$, so

$$SP_2(y^1,w)=SP_2(y^2,w)+\frac{2}{n\bar{y}^1}\left(y_j^1-y_i^1\right)\left\{-\frac{4}{n-1}\left[\frac{(n-1)\left(\frac{n+1}{2}\right)-(n-1)}{4n}-\frac{2i-1}{2n}\right]\right.$$
$$\left.+\frac{n-5}{n-1}\frac{2i-1}{2n}+\frac{4}{n+1}\left[\frac{(n+1)\left(\frac{3n+1}{2}\right)-(n+1)}{4n}-\frac{2j-1}{2n}\right]-\frac{n-3}{n+1}\frac{2j-1}{2n}\right\}$$

$$SP_2(y^1,w)=SP_2(y^2,w)+\frac{2}{n^2\bar{y}^1}\left(y_j^1-y_i^1\right)(n+i-j).$$

Since $y_j^1-y_i^1>0$ and $i-j>-n$, then $\frac{2}{n^2\bar{y}^1}\left(y_j^1-y_i^1\right)(n+i-j)>0$, and so $SP_2(y^1, w)>SP_2(y^2, w)$. Thus $SP_2$ satisfies the increased spread axiom.

The proof for $SP_1$ is very similar and omitted.

Proof that $SP_1$ and $SP_2$ satisfy the increased bipolarity axiom:

If $(y^{2-}, w^-)$ is derived from $(y^{1-}, w^-)$ via a sequence of strict equalizing transfers, or if $(y^{2+}, w^+)$ is derived from $(y^{1+}, w^+)$ via a sequence of strict equalizing transfers, then $C^W(y^2, w)<C^W(y^1, w)$. Since $C^B(y^2, w)=C^B(y^1, w)$, $\bar{y}^2=\bar{y}^1$ and $m_{y^2}=m_{y^1}$, then using equations (6) and (8) we get $SP_1(y^2, w)>SP_1(y^1, w)$ and $SP_2(y^2, w)>SP_2(y^1, w)$. Thus $SP_1$ and $SP_2$ satisfy the increased bipolarity axiom.

**Proof 2.**
OLS yield

$$\hat{\gamma}=\frac{\sum_i R_{w,i}.2\sigma^2_{R_w}\frac{2y_i^M-y_i}{\bar{y}}-n.mean\left(R_{w,i}\right).mean\left(2\sigma^2_{R_w}\frac{2y_i^M-y_i}{\bar{y}}\right)}{n\sigma^2_{R_w}}. \quad (16)$$

Now $mean\left(R_{w,i}\right)=\frac{1}{2}$ and $mean\left(2\sigma^2_{R_w}\frac{2y_i^M-y_i}{\bar{y}}\right)=2\sigma^2_{R_w}$.

We replace these means by their expression in (16) and we get:

$$\hat{\gamma}=\frac{2}{n\bar{y}}\sum_{i=1}^{n}\left(2y_i^M-y_i\right)R_{w,i}-1.$$

**Proof 3.**
For $n$ odd, the cumulative proportion of health of the 50 percent of individuals whose income is the lowest is by definition the following:

$$L_{y,w}(0.5) = \frac{1}{n\bar{y}} \sum_{i=1}^{(n+1)/2} y_i.$$

Since $y_i = \alpha + \sum_k \beta_{x_k} x_{ki} + \varepsilon_i$, we get:

$$L_{y,w}(0.5) = \frac{\alpha}{2\bar{y}} \frac{n+1}{n} + \sum_k \frac{\beta_{x_k}}{\bar{y}} \frac{1}{n} \sum_{i=1}^{(n+1)/2} x_{ki} + \frac{1}{n\bar{y}} \sum_{i=1}^{(n+1)/2} \varepsilon_i.$$

For $n$ large enough, $\frac{n+1}{n} \approx 1$, and so:

$$L_{y,w}(0.5) \approx \frac{\alpha}{2\bar{y}} + \sum_k \frac{\beta_{x_k} \bar{x}_k}{\bar{y}} \frac{1}{n\bar{x}_k} \sum_{i=1}^{(n+1)/2} x_{ki} + \frac{1}{n\bar{y}} \sum_{i=1}^{(n+1)/2} \varepsilon_i$$
$$\Leftrightarrow L_{y,w}(0.5) \approx \frac{\alpha}{2\bar{y}} + \sum_k \frac{\beta_{x_k} \bar{x}_k}{\bar{y}} L_{x_k,w}(0.5) + \frac{1}{n\bar{y}} \sum_{i=1}^{(n+1)/2} \varepsilon_i$$

where $L_{x_k,w}(0.5)$ is the concentration curve for factor $x_k$ evaluated at the 50th percentile. So,

$$0.5 - L_{y,w}(0.5) \approx 0.5 - \frac{\alpha}{2\bar{y}} - \sum_k \frac{\beta_{x_k} \bar{x}_k}{\bar{y}} L_{x_k,w}(0.5) - \frac{1}{n\bar{y}} \sum_{i=1}^{(n+1)/2} \varepsilon_i.$$

Bearing in mind that:

$$\bar{y} = \alpha + \sum_k \beta_{x_k} \bar{x}_k$$
$$\Leftrightarrow \alpha = \bar{y} - \sum_k \beta_{x_k} \bar{x}_k$$
$$\Leftrightarrow \frac{\alpha}{2\bar{y}} = 0.5 - \sum_k \frac{\beta_{x_k} \bar{x}_k}{\bar{y}} 0.5$$

we obtain:

$$0.5 - L_{y,w}(0.5) \approx \sum_k \frac{\beta_{x_k} \bar{x}_k}{\bar{y}} \left(0.5 - L_{x_k,w}(0.5)\right) - \frac{1}{n\bar{y}} \sum_{i=1}^{(n+1)/2} \varepsilon_i.$$

Moreover equation (1) indicates that $C(y,w) = \sum_k \frac{\beta_{x_k} \bar{x}_k}{\bar{y}} C(x_k, w) + \frac{GC_\varepsilon}{\bar{y}}$.

- We replace $0.5 - L_{y,w}(0.5)$ and $C(y, w)$ by their expressions in equation (3):

$$SP_1(y,w) \approx \frac{2}{m_y}\sum_k \beta_{x_k}\overline{x}_k\big(2\big(0.5 - L_{x_k,w}(0.5)\big) - C(x_k,w)\big) + \frac{2\overline{y}}{m_y}GSP_\varepsilon$$

with $GSP_\varepsilon = -\frac{2}{n\overline{y}}\sum_{i=1}^{(n+1)/2}\varepsilon_i - \frac{GC_\varepsilon}{\overline{y}}$

$$SP_1(y,w) \approx \sum_k \frac{\beta_{x_k} m_{x_k}}{m_y}\frac{2\overline{x}_k}{m_{x_k}}\big[2\big(0.5 - L_{x_k,w}(0.5)\big) - C(x_k,w)\big] + \frac{2\overline{y}}{m_y}GSP_\varepsilon.$$

Since the polarization index for $x_k$ is:

$$SP_1(x_k,w) = \frac{2\overline{x}_k}{m_{x_k}}\big(2\big(0.5 - L_{x_k,w}(0.5)\big) - C(x_x,w)\big)$$

we obtain $SP_1(y,w) \approx \sum_k \frac{\beta_{x_k} m_{x_k}}{m_y} SP_1(x_k,w) + GSP_\varepsilon$ .

- We replace $0.5 - L_{y,w}(0.5)$ and $C(y, w)$ by their expressions in equation (4):

$$SP_2(y,w) \approx \sum_k \frac{\beta_{x_k}\overline{x}_k}{\overline{y}}\big(2\big(0.5 - L_{x_k,w}(0.5)\big) - C(x_k,w)\big) + GSP_\varepsilon$$

with $GSP_\varepsilon = -\frac{2}{n\overline{y}}\sum_{i=1}^{(n+1)/2}\varepsilon_i - \frac{GC_\varepsilon}{\overline{y}}$ .

Since the polarization index for $x_k$ is:

$$SP_2(x_k,w) = 2\big(0.5 - L_{x_k,w}(0.5)\big) - C(x_k,w)$$

we obtain $SP_2(y,w) \approx \sum_k \frac{\beta_{x_k}\overline{x}_k}{\overline{y}} SP_2(x_k,w) + GSP_\varepsilon$ .

Note: If $n$ is even,

$$L_{y,w}(0.5) = \frac{1}{n\overline{y}}\sum_{i=1}^{n/2} y_i$$

$$L_{y,w}(0.5) = \frac{\alpha}{2\overline{y}} + \sum_k \frac{\beta_{x_k}}{\overline{y}}\frac{1}{n}\sum_{i=1}^{n/2} x_{ki} + \frac{1}{n\overline{y}}\sum_{i=1}^{n/2}\varepsilon_i$$

Thus $SP_1(y,w) = \sum_k \frac{\beta_{x_k} m_{x_k}}{m_y} SP_1(x_k,w) + GSP_\varepsilon$

and $SP_2(y,w) = \sum_k \frac{\beta_{x_k}\overline{x}_k}{\overline{y}} SP_2(x_k,w) + GSP_\varepsilon$ .

## References


Abul Naga, Ramses H., and Pierre-Yves Geoffard, "Decomposition of Bivariate Inequality Indices by Attributes," *Economics Letters*, 90, 362–7, 2006.

Atkinson, Anthony B. and Francois Bourguignon, "The Comparison of Multi-Dimensioned Distributions of Economic Status," *Review of Economic Studies*, XLIX, 183–201, 1982.

Boland, Philip J. and Frank Proschan, "Multivariate Arrangement Increasing Functions with Application in Probability and Statistics," *Journal of Multivariate Analysis*, 25, 286–98, 1988.

Cutler, David M., Elizabeth Richardson, Theodore E. Keeler, and Douglas Staigers, "Measuring the Health of the U.S. Population," *Brookings Papers on Economic Activity. Microeconomics*, 217–82, 1997.

Deaton, Angus S. and Christina Paxson, "Aging and Inequality in Income and Health," *The American Economic Review, Papers and Proceedings*, 88, 1998.

Deutsch, Jacques, Meital Hanoka, and Jacques Silber, "On the Link Between the Concepts of Kurtosis and Bipolarization," *Economics Bulletin*, 4, 1–5, 2007.

Groot, Wim, "Adaptation and Scale of Reference Bias in Self-Assessments of Quality of Life," *Journal of Health Economics*, 19, 403–20, 2000.

Idler, Ellen L. and Yoav Benyamini, "Self-Rated Health and Mortality: A Review of Twenty-Seven Community Studies," *Journal of Health and Social Behavior*, 38, 21–37, 1997.

Kakwani, Nanak C., Adam Wagstaff, and Eddy Van Doorslaer, "Socioeconomic Inequalities in Health: Measurement, Computation and Statistical Inference," *Journal of Econometrics*, 77, 87–103, 1997.

Oaxaca, Ronald, "Male–Female Wage Differentials in Urban Labor Markets," *International Economic Review*, 14, 693–709, 1973.

O'Donnell, Owen, Eddy Van Doorslaer, Adam Wagstaff, and Magnus Lindelow, *Analyzing Health Equity using Household Survey Data: A Guide to Techniques and Their Implementation*, The World Bank, 2006.

Rodríguez, Juan and Rafael Salas, "Extended Bi-Polarization and Inequality Measures," *Research on Economic Inequality*, 9, 69–83, 2003.

Tsui, Kai-Yuen, "Multidimensional Inequality and Multidimensional Generalized Entropy Measures: An Axiomatic Derivation," *Social Choice and Welfare*, 16, 145–57, 1999.

Van Doorslaer, Eddy and Andrew M. Jones, "Inequalities in Self-Reported Health: Validation of a New Approach to Measurement," *Journal of Health Economics*, 22, 61–78, 2003.

Van Doorslaer, Eddy and Xander Koolman, "Explaining the Differences in Income-Related Health Inequalities Across European Countries," *Health Economics*, 13, 609–28, 2004.

Wagstaff, Adam and Eddy Van Doorslaer, "Overall Versus Socioeconomic Health Inequality: A Measurement Framework and Two Empirical Illustrations," *Health Economics*, 13, 297–301, 2004.

Wagstaff, Adam, Eddy Van Doorslaer, and Pierella Paci, "Equity in the Finance and Delivery of Health Care: Some Tentative Cross-Country Comparison," *Oxford Review of Economic Policy*, 5, 89–112, 1989.

Wagstaff, Adam, Pierella Paci, and Eddy Van Doorslaer, "On the Measurement of Inequalities in Health," *Social Sciences and Medecine*, 33, 545–57, 1991.

Wagstaff, Adam, Eddy Van Doorslaer, and Naoko Watanabe, "On Decomposing the Causes of Health Sector Inequalities with an Application to Malnutrition Inequalities in Vietnam," *Journal of Econometrics*, 112, 207–23, 2003.

Wang, You-Qiang and Kai-Yuen Tsui, "Polarization Orderings and New Classes of Polarization Indices," *Journal of Public Economic Theory*, 2, 349–63, 2000.

Wolfson, Michael C., "When Inequalities Diverge," *American Economic Review*, 84, 353–8, 1994.