
WORKING PAPERS

Disentangling the impacts of circumstances and efforts on health inequality: the case of Luxembourg

Joseph **DEUTSCH** ¹
Maria Noel **PI ALPERIN** ²
Jacques **SILBER** ^{1,2}

¹ Bar-Ilan University, Israel

² LISER, Luxembourg

LISER Working Papers are intended to make research findings available and stimulate comments and discussion. They have been approved for circulation but are to be considered preliminary. They have not been edited and have not been subject to any peer review.

*The views expressed in this paper are those of the author(s) and do not necessarily reflect views of LISER.
Errors and omissions are the sole responsibility of the author(s).*

Disentangling the impacts of circumstances and efforts on health inequality: the case of Luxembourg¹

Joseph Deutsch

Department of Economics, Bar-Ilan University, 52900 Ramat-Gan, Israel.

E-mail: yosef.deutsch@gmail.com

María Noel Pi Alperin

LISER, 11 Porte des Sciences, L-4366 Esch/Alzette, Luxembourg.

E-mail: MariaNoel.PiAlperin@liser.lu

Jacques Silber

Department of Economics, Bar-Ilan University, Israel and Senior Research Fellow,
LISER, Esch/Alzette, Luxembourg.

E-mail: jsilber_2000@yahoo.com.

May 2016

Abstract: This paper attempts to determine the contribution of circumstances, efforts (and lifestyle) and demographic variables (age and gender) to inequality in health in Luxembourg. Health is measured subjectively by self-assessed health and is considered first as a binary variable, then as an ordinal variable. The educational level of each parent, the financial situation of the family during childhood and the area of birth are considered as circumstances while effort and lifestyle variables are proxied by information on the educational level of the individual, whether he/she smoked and whether he/she had a physical activity on a regular basis. The respective impacts of the three categories of explanatory variables (circumstances, effort and demographic variables) on health inequality are derived via a Shapley decomposition of the pseudo R-square of logit regressions. Differences in circumstances and effort and lifestyle explain each around a quarter of the pseudo R-square.

Key Words: circumstances – efforts - equality of opportunity – health – Luxembourg - Shapley decomposition

¹ Acknowledgement: This research is part of the HEADYNAP project supported by the Luxembourg 'Fonds National de la Recherche' (contract FNR C12/SC/3977324/HEADYNAP) and by core funding for LISER from the Ministry of Higher Education and Research of Luxembourg.

1. Introduction

Equality of opportunity is not a simple notion and there is no consensus on the way it should be defined. Even the notion of "leveling the playing field" which has often been identified with that of equality of opportunity may be understood differently, depending on the philosophical position one takes.

Among economists the most thorough proposal of implementing the concept of equality of opportunity has probably been made by Roemer (see, in particular, Roemer, 1998) who summarized his ideas as follows in the *Boston Review* (Roemer, 1995). Stressing that many people make a link between egalitarianism and the welfare state whose implementation tends often to absolve the individuals from any responsibility, Roemer argues that society should provide a "level playing field" and once this is done individuals should be responsible for the consequences of their own choices. The question is then to find out what is required to level the playing field. A first approach requires the government to make sure legal barriers to social mobility are eliminated so that, for example, companies will hire workers or students will be accepted in educational institutions according to merit. As far as education is concerned, a second approach identifies equal opportunity with a guarantee of equal access to education for everyone, which implies that individuals will all receive equal amounts of the various resources provided by society for educational purposes. A third approach suggested by Roemer, and emphasizing the concept of personal responsibility, argues that equality-of-opportunity may well imply that people will receive quite unequal amounts of the relevant resources. Roemer mentions, for example, the case of parents who have several children, one of them with learning problems. For Roemer equality of opportunity requires that parents spend more money on this child and eventually also more time on him/her. Note however that equality of opportunity does not ask society "to insure individuals against bad results, when they are the consequences of individual choices made after opportunities have been equalized" (Roemer, 1995).

A distinction should also be made between *brute luck* (see, Dworkin, 2000) which happens when, for example, you are hit by a car that runs a red light and you walked in the pedestrian crossing, and *option luck* which is the case where the car hits you while you are "jay

walking." An equal opportunity approach might well recommend covering only the risk related to "brute luck".

More generally, Roemer makes a distinction between *circumstances*, which are beyond an individual's control, and *autonomous choices* which are within his/her control. As a consequence, Roemer recommends that society compensates only the cases where bad consequences are due to circumstances or brute luck. There is no need to offer an insurance against the implications of an individual's autonomous choices. In Roemer's (1995) words "an equal opportunity policy must equalize outcomes in so far as they are the consequences of causes beyond a person's control, but allow differential outcomes in so far as they result from autonomous choice". Naturally, this distinction requires us to be able to know which type of an individual's behavior is due to circumstances and which one results from autonomous choice. The Roemer approach makes therefore a distinction between types (that is, factors like innate abilities or disabilities that should be compensated for) and what he generally calls *effort* (that is, factors like preferences or ambition which are under the responsibility of the individual) and then recommends to equalize, for a given level of *effort*, differences in individual outcomes which are the consequence of differences in *types*.

In addition to making a distinction between the concepts of circumstances and efforts the literature on equality of opportunity stresses also the notion of luck. Lefranc *et al.* (2009) mention in fact four rather than two types of luck. They call the first one *social background luck*. This is a factor over which the individual has no control and hence it should be classified as what was defined previously as circumstance. Note that this factor corresponds to the concept of *social lottery* developed by Rawls. The second category is what they called *genetic luck* (e.g. a given talent), a factor which is also beyond an individual's control. This factor corresponds to what Rawls had defined as *natural lottery*. To illustrate the third type of luck, Lefranc *et al.* (2006) take the case of twin brothers in the late 1960s or early 1970s in the United States who have similar talents and the same social background. One of them however, as a result of the draft lottery, had to go to Vietnam and when he came back he did not study and had a low income over his lifetime. This situation corresponds to what Dworkin has called *brute luck*. The fourth category of luck

refers to the case where two individuals have a choice between two lotteries. The outcome of the first lottery is certain while that of the second is random. If we assume that the individuals made different choices and ended up with different outcomes this result will be the consequence of what Dworkin called *option luck*, but Lefranc *et al.* (2006) emphasize the fact that what is assumed here is that it is really *informed option luck*.

This framework of analysis emphasizing the notions of circumstances, efforts and luck has by now been applied in many papers devoted to the measurement of inequality of opportunity in general, inequality of opportunity in health in particular. The focus of the present paper is also on this topic but our approach is innovative because the respective impacts of circumstances and efforts on health inequality, as well as that of demographic variables, are estimated by implementing a Shapley decomposition of the Pseudo R-square of a health logit regression.

We illustrate our approach by computing the contribution of circumstances (measured as the educational level of parents, the economic situation of the family during childhood, parents' and individuals' country of birth, and the years of immigration), the efforts (measured as the individuals' health related behavior with respect to smoking and to physical activity, and the educational achievement) and the demographic variables (such as gender and age) on inequality in health using Luxembourgish data.

The paper is organized as follows. Section 2 takes a look at the studies that attempt to measure inequality of opportunity in health. Section 3 briefly describes the Panel Socio-Economique *Liewen zu Lëtzebuerg* (PSELL-3) and presents the three sets of variables used to characterize the notions of health, circumstances and efforts. Section 4 describes in details the methodology implemented in the present paper while Section 5 gives an empirical illustration based on Luxembourgish data. Concluding comments will be presented in Section 6.

2. On Inequality of Opportunity in Health

Motivated by the concept of inequality of opportunity developed by Dworkin (1981), Arneson (1989), Cohen (1989), Roemer (1998), and Fleurbaey (2008), a growing number of articles try to distinguish between the legitimate and the illegitimate causes of health inequalities.

In order to measure equality of opportunity in health the notion of *effort* has to be complemented by that of *lifestyle* which is assumed to be at least partly the consequence of personal choice. Such a lifestyle may, for example, include the choice of diet or the decision to have regularly some physical exercise. Fleurbaey and Schokkaert (2009) take in fact such an approach in their study of inequalities in health and health care. They thus first define two conditions. The first one states that a measure of unfair health inequality should not reflect legitimate variation in outcomes, that is, inequalities which are caused by differences in the responsibility variables. According to the second condition if a measure of unfair inequality is zero, there should be no illegitimate differences left, that is, two individuals with the same value for the responsibility variable should have the same outcome. As is by now well-known these two conditions are incompatible.² Fleurbaey and Schokkaert (2009) propose therefore two methods to measure unfair health inequalities. In the first case they assume that the health h_i of an individual depends on his/her income y_i and lifestyle l_i . Their first approach consists of constructing a hypothetical distribution of individual health $\tilde{h}_i(y_i, \tilde{l})$ where individual health would be a function of income and some reference lifestyle $\tilde{l}$. The inequality of such a distribution is called *direct unfairness* by the authors. Such an inequality clearly reflects only variation due to differences in income (circumstances), since differences in lifestyle have been eliminated. In the second case a distribution is constructed where all the illegitimate sources of variation have been removed by fixing a value y^* for income and defining a reference health level h_i^* as $h(y^*, l_i)$. One would then compute the inequality of the vector $(h_i - h_i^*)$ and such an inequality corresponds to what the authors called “fairness gap”.

² See, Fleurbaey (2008) for more details on this issue.

While Fleurbaey and Schokkaert's (2009) paper was only methodological, there have been some empirical attempts to measure inequality of opportunity in health. Rosa Dias (2009), for example, used data from the UK *National Child Development Study*. Circumstances were proxied by parental socioeconomic status and childhood health. More precisely the circumstance variables were the weight at birth, dummy variables for whether the mother smoked after the fourth month of pregnancy and for whether the child was breastfed, a set of morbidities experienced by the child up until the age of 16, dummy variables for the case where parents had chronic diseases and for the incidence of hereditary conditions such as diabetes and epilepsy among parents, brothers and sisters and dummy variables for whether the child was obese at age 16 and for whether both parents were smokers in 1974. Effort was measured via information on health-related lifestyles such as cigarette smoking, alcohol consumption, consumption of fried food and educational attainment. Clearly these variables are likely to be constrained by circumstances but they also reflect individual choices. Note also that all the variables used to proxy lifestyles were based on self-reported information. Finally, the health outcome was apprehended via self-assessed health measured on a four points scale: excellent, good, fair and poor health. Rosa Dias (2009) implemented then stochastic dominance tests to detect inequality of opportunity in the conditional distributions of self-assessed health in adulthood.

In another paper, Rosa Dias (2010) takes unobserved heterogeneity into account to measure inequality of opportunity in health. This allows him to address the "partial-circumstance problem", that is, the fact that some circumstances are not observed. In addition, Rosa Dias (2010) extends the examination of inequality of opportunity to health outcomes other than self-assessed health, such as long-standing illness, disability and mental health.

Trannoy *et al.* (2010) attempt also to estimate the degree of inequality in health opportunity but the focus of their analysis is on older adults because their database is the French part of SHARE (*Survey of Health, Ageing and Retirement in Europe*), a survey conducted in many European countries but limited to individuals who are at least 50 years old. In order to neutralize the impact of circumstances on efforts, the authors regress the two effort variables, the educational level and the social status, in two separated equations against the vector of circumstances. They then introduce the estimated residuals of these two equations

into the third equation explaining health in adulthood along with the vector of circumstances. The impact of circumstances and efforts on inequality in health opportunity is then measured via the Gini index.

Lazar (2013), examining inequality in health opportunity in Israel, adopts an approach very similar to that of Trannoy *et al.* (2010) but she combines it with that used by Checchi and Peragine (2010) for the case of earnings, making thus a distinction between ex-ante and ex-post inequality in health opportunity. More precisely, Checchi and Peragine (2010) explain that in an ex-ante approach, whose focus is on circumstances (types), there will be equality of opportunity if all the types have the same mean income. In an ex-post approach which emphasizes the concept of “tranches” (a tranche corresponds to a given level of effort) there will be equality of opportunity if all the individuals who exert the same level of effort have the same outcome.

All these papers measure the impact of individuals’ efforts and childhood circumstances in adulthood health without having a special look at the way these two sets of variables are correlated. In Roemer’s definition of equality of opportunity, the individuals’ efforts must be purged from childhood circumstances (Roemer, 1998), while following Barry (2005) the individuals’ efforts have to be fully respected whatever the influence of past circumstances on efforts will be. Finally, according to Swift (2005) the parents’ own effort have to be fully respected in order to encourage parents to transmit a value on efforts to their descendants. In this context, the papers of Jusot *et al.* (2013) and Bricard *et al.* (2013) assess whether it empirically matters which normative way of treating the correlation between circumstances and efforts is adopted in the measurement of inequalities of opportunity in health. In particular, Jusot *et al.* (2013) adopt these three alternative normative ways in their paper and provide an empirical evaluation in France. They show that the share of inequalities of opportunities does not vary much according to the adopted definition. In their paper, Bricard *et al.* (2013) compare inequalities of opportunity in health in Europe using SHARE data. They estimate whether it is empirically important to adopt Barry or Roemer viewpoint. They found that among the 13 countries analyzed (two Scandinavian countries, six countries from Western Europe, three Mediterranean countries and two transition countries) it was not possible to find a general pattern on the relationship

between the extent of inequalities of opportunities and the way the correlation between efforts and circumstances matters for the assessment of inequalities of opportunity.

In our paper, we assume that an individual's health depends first on age and gender, variables which we label *demographic variables*, second on circumstances which are beyond his/her responsibility and third on characteristics which belong to the sphere of his/her responsibility. It is however likely that some of these individual characteristics which may be considered as *efforts* are correlated with an individual's circumstances. Then, there is first a direct effect of circumstances measuring the impact of various circumstances on individual health, second an indirect impact which acts via the influence of some circumstances on the efforts of the individual. In making such a distinction we follow in fact Trannoy *et al.* (2010) and Lazar (2013) who, as mentioned previously, recommended a procedure in two-stages.

3. The data sources

The Panel Socio-Economique *Liewen zu Lëtzebuerg* (PSELL-3), is a general purpose panel survey carried out annually since 2003 with an initial sample of over 3500 households representative of the population living in private dwellings in Luxembourg. This panel has comparable variables over the different waves but includes also specific and different questions in each wave. The analysis of this paper is mainly based on data for the year 2008 (wave 6), but we also use some information from waves 3 and 5 (2005 and 2007, respectively). In particular, wave 3 and 5 include information on the educational level and the country of birth of an individual's parents. Hence, we matched this information about parents from previous waves with the information on the individual given by wave 6. This matching is possible since the information matched is exactly the same over the various years.

We restrict our sample to residents who are 25 to 65 years old. The main reason is that only for this interval of age we have individuals who answered all the selected questions, especially the questions about their parents. After excluding all observations with missing

values on any of the variables used in this study, our final sample includes 2332 individual observations.

Concerning the variables used in this paper, the list was selected on the basis of what was available in the survey. In particular, in our empirical analysis, the individual health status is viewed as a function of three main sets of variables: the efforts as measured by lifestyles, the circumstances as measured by family background and the demographic characteristics including gender and age. The descriptive statistics of the main variables used in this paper are presented in Table 1.

3.1. The health variable

One can find in the literature several ways of measuring health though each refers to one of the three dimensions of an individual's health status: subjective health (self-assessed health, symptoms and quality of life); medical health (diagnosed or reported diseases); or functional health (functional limitations) (Sermet and Cambois, 2002; Blaxter, 1985).

In this paper we use a subjective measure, the self-perception health, which is one of the most common collected measurements of health in surveys (Tubeuf *et al.*, 2008). This variable turns out to be a very good predictor of mortality (Idler and Benyamini, 1997), of health care utilization (DeSalvo *et al.*, 2005), and can be multidimensional (it aggregates the perception about different items of health).

In *PSELL-3*, individuals have been asked: "In general would you say that your health is ... very good, good, fair, poor or very poor?" Based on this question, we have then considered health as an ordinal variable with a five points scale, and as a binary variable grouping very good and good health on one side and fair, poor and very poor on the other side. In our database, 76% of individuals in Luxembourg reported that their health was good or very good.

3.2. The effort variables

Given the information available in the *PSELL-3* survey for 2008, we have included two categories of variables reflecting efforts. The first one is lifestyle measured by the individuals' health related behavior with respect first to smoking (variable equal to 1 if the individual does not smoke and to 0 otherwise), second to physical activity (variable equal to 1 if he/she has such an activity and to 0 otherwise).

The other effort category is educational achievement measured as the highest educational level of individuals (variable equal to 1 if the individual studied beyond high school and to 0 otherwise).

Concerning the lifestyles variables, whereas 76% of respondents were currently non-smokers, only 34% of the individuals reported that they have a physical activity on a regular basis. Finally, only 30% of the individuals studied beyond high school.

3.3. The circumstance variables

Seven explanatory variables were assumed to be part of the circumstances. First, there is the educational level of each of the parents.³ Here a distinction was made between two categories: those who have either no education, a primary or secondary education, and those who studied beyond high school (this variable is equal to 1 if the father, or the mother, studied beyond high school and to 0 otherwise). Only 7% of the mothers and 17% of the fathers have a post-secondary educational level.

Another circumstance variable was the economic situation of the family during childhood. Individuals were thus asked whether their family used to have financial difficulties when they were between 12 and 16 years old. This variable, labeled *financial*, was assumed to take two values: for those who never, or occasionally, had economic problems the variable was equal to 1, while for those families who often, or most of the time, had financial

³ For those individuals stating that they do not have a father or a mother, the variable was assumed to be equal to 0.

difficulties the variable was equal to 0. Specifically, 78% of the individuals declared that they never had financial difficulties during childhood.

Luxembourg is the country with the largest share of immigrants in the European Union (European Commission, 2011). 41% of foreign immigrants are Portuguese, constituting the largest foreign community in the Grand Duché (Berger, 2008). This is the main reason why the ‘*parents’ country of birth*’, ‘*the individuals’ country of birth*’ and the ‘*year of immigration*’⁴ were considered as circumstance variables having eventually an incidence on an individual’s health status. For the individuals as well as their parents, four dummy variables referring to the area of birth have been considered: Luxembourg, Portugal, Other EU-15 countries (Portugal and Luxembourg excluded), and other countries.⁵ Concerning the individuals who answered the survey, 61% of them are natives and 11% Portuguese. 12% is the percentage of mothers and fathers coming from Portugal and 44% and 46% respectively from Luxembourg.

Finally, the third set of explanatory variables of the health status is composed of demographic variables. In particular, we included gender (variable equal to 1 for male individuals) and age.⁶ Following Jusot *et al.* (2013), these are considered as biological determinants of health status.

4. Inequality in health opportunity in Luxembourg: the methodology adopted

Let φ_i be the health level of individual i , the variable *health*, previously defined, is a subjective measure of self-assessed health. In a first stage of the analysis, this variable will be equal to 1 if health is very good or good and to 0 otherwise (precisely, if health is fair, poor or very poor).⁷ c_{ki} denotes the level of the circumstance variable k ($k = 1, \dots, K$) for

⁴ *Year of immigration* is a variable defined as being equal to the difference between 2008 and the year the individual arrived in Luxembourg.

⁵ For technical reasons, a variable *Parents born in Luxembourg* has been constructed. The variable is equal to 1 if at least one of the parents is/was an immigrant.

⁶ We have also included its square as explanatory variable, *Square of Age*.

⁷ In a second stage, the health variable will be considered as an ordinal variable on a five point scale.

individual i and f_{hi} ($h = 1, \dots, H$) the effort level of individual i . Finally, let us call d_{li} the value of the demographic variable l ($l = 1, \dots, L$) for individual i .

When the health variable φ_i is binary, we can then write that,

$$\begin{aligned} Prob\{\varphi_i = 1\} &= F\left(\sum_{k=1}^K \beta_k c_{ki} + \sum_{h=1}^H \gamma_h f_{hi} + \sum_{l=1}^L \delta_l d_{li} + u_i\right) \\ &= \left(e^{\sum_{k=1}^K \beta_k c_{ki} + \sum_{h=1}^H \gamma_h f_{hi} + \sum_{l=1}^L \delta_l d_{li} + u_i}\right) \cdot \left(1 + e^{\sum_{k=1}^K \beta_k c_{ki} + \sum_{h=1}^H \gamma_h f_{hi} + \sum_{l=1}^L \delta_l d_{li} + u_i}\right)^{-1}, \end{aligned} \quad (1)$$

and

$$\begin{aligned} Prob\{\varphi_i = 0\} &= 1 - F\left(\sum_{k=1}^K \beta_k c_{ki} + \sum_{h=1}^H \gamma_h f_{hi} + \sum_{l=1}^L \delta_l d_{li} + u_i\right) \\ &= \left(1 + e^{\sum_{k=1}^K \beta_k c_{ki} + \sum_{h=1}^H \gamma_h f_{hi} + \sum_{l=1}^L \delta_l d_{li} + u_i}\right)^{-1}, \end{aligned} \quad (2)$$

where u_i is the residual⁸ of the regression for individual i . Let v_i , w_i and z_i be defined as,

$$v_i = \sum_{k=1}^K \beta_k c_{ki}. \quad (3)$$

$$w_i = \sum_{h=1}^H \gamma_h f_{hi}. \quad (4)$$

$$z_i = \sum_{l=1}^L \delta_l d_{li}. \quad (5)$$

Combining expressions (1) to (5) we derive that,

$$Prob\{\varphi_i = 1\} = (e^{v_i + w_i + z_i + u_i}) \cdot (1 + e^{v_i + w_i + z_i + u_i})^{-1}, \quad (6)$$

and

$$Prob\{\varphi_i = 0\} = (1 + e^{v_i + w_i + z_i + u_i})^{-1}. \quad (7)$$

To have an idea of the goodness-of-fit of the logit regressions we use a criterion that is similar to the R -square used in linear regressions. The idea is to compute the maximal value of the log-likelihood ($\ln L$) and compare it with the log-likelihood obtained when only a constant term is introduced ($\ln L_0$). The likelihood ratio, LRI , is then defined as,

$$LRI = 1 - \frac{\ln L}{\ln L_0}, \quad (8)$$

⁸ Luck is one of the factors affecting the residual. On the impact of luck on equality of opportunity, see, Lefranc *et al.* (2006).

where the bounds of this measure are 0 and 1 (Greene, 1992, pp. 651–653).

The Shapley decomposition

To determine the impact of the different categories of explanatory variables we use the concept of Shapley decomposition (see, Chantreuil and Trannoy, 2013; Shorrocks, 2013, and Sastre and Trannoy, 2002) which is by now quite well known. We apply this decomposition to determine the respective impact of circumstances, efforts and demographic variables on the likelihood ratio LRI . In other words, we take into account all possible combinations of circumstances, efforts and demographic variables, via the so-called Shapley decomposition procedure.

The actual likelihood ratio can be written as,

$$LRI_1 = LRI(c_{ki} \neq 0; f_{hi} \neq 0; d_{li} \neq 0).$$

Assume, for example, that we do not include the circumstances, c_{ki} , of the different individuals in the regression. In such a case the likelihood ratio will be expressed as,

$$LRI_2 = LRI(c_{ki} = 0; f_{hi} \neq 0; d_{li} \neq 0).$$

Similarly, assume that we do not include in the regression the effort variables, f_{hi} . In such a case we will define the likelihood ratio as,

$$LRI_3 = LRI(c_{ki} \neq 0; f_{hi} = 0; d_{li} \neq 0).$$

We can also assume that we do not introduce in the regression the demographic variables, d_{li} , in which case the likelihood ratio will be,

$$LRI_4 = LRI(c_{ki} \neq 0; f_{hi} \neq 0; d_{li} = 0).$$

Naturally, we could also decide not to include two sets of explanatory variables (e.g. the circumstance and the effort variables, the circumstances and the demographic variables, the effort and the demographic variables, named respectively, LRI_5 , LRI_6 , and LRI_7).

Using the by now well-known Shapley procedure and derive that the contribution of circumstances, C_{circ} , to the overall actual likelihood ratio, LRI_1 , may be expressed as,

$$\begin{aligned}
C_{circ} = & \left(\frac{2}{6}\right) \{LRI(c_{ki} \neq 0; f_{hi} \neq 0; d_{li} \neq 0) - LRI(c_{ki} = 0; f_{hi} \neq 0; d_{li} \neq 0)\} \\
& + \left(\frac{1}{6}\right) \{LRI(c_{ki} \neq 0; f_{hi} \neq 0; d_{li} = 0) - LRI(c_{ki} = 0; f_{hi} \neq 0; d_{li} = 0)\} \\
& + \left(\frac{1}{6}\right) \{LRI(c_{ki} \neq 0; f_{hi} = 0; d_{li} \neq 0) - LRI(c_{ki} = 0; f_{hi} = 0; d_{li} \neq 0)\} \\
& + \left(\frac{2}{6}\right) \{LRI(c_{ki} \neq 0; f_{hi} = 0; d_{li} = 0) - LRI(c_{ki} = 0; f_{hi} = 0; d_{li} = 0)\}
\end{aligned}$$

$$C_{circ} = \left(\frac{2}{6}\right) (LRI_1 - LRI_2) + \left(\frac{1}{6}\right) (LRI_4 - LRI_6) + \left(\frac{1}{6}\right) (LRI_3 - LRI_5) + \left(\frac{2}{6}\right) (LRI_7), \quad (9)$$

since by definition, $LRI(c_{ki} = 0; f_{hi} = 0; d_{li} = 0) = 0$.

Similarly, the contribution of efforts, C_{eff} , to the actual likelihood ratio, LRI_1 , may be expressed as,

$$\begin{aligned}
C_{eff} = & \left(\frac{2}{6}\right) \{LRI(c_{ki} \neq 0; f_{hi} \neq 0; d_{li} \neq 0) - LRI(c_{ki} \neq 0; f_{hi} = 0; d_{li} \neq 0)\} \\
& + \left(\frac{1}{6}\right) \{LRI(c_{ki} \neq 0; f_{hi} \neq 0; d_{li} = 0) - LRI(c_{ki} \neq 0; f_{hi} = 0; d_{li} = 0)\} \\
& + \left(\frac{1}{6}\right) \{LRI(c_{ki} = 0; f_{hi} \neq 0; d_{li} \neq 0) - LRI(c_{ki} = 0; f_{hi} = 0; d_{li} \neq 0)\} \\
& + \left(\frac{2}{6}\right) \{LRI(c_{ki} = 0; f_{hi} \neq 0; d_{li} = 0) - LRI(c_{ki} = 0; f_{hi} = 0; d_{li} = 0)\}
\end{aligned}$$

$$C_{eff} = \left(\frac{2}{6}\right) (LRI_1 - LRI_3) + \left(\frac{1}{6}\right) (LRI_4 - LRI_7) + \left(\frac{1}{6}\right) (LRI_2 - LRI_5) + \left(\frac{2}{6}\right) (LRI_6). \quad (10)$$

Finally, the contribution of the demographic variables, C_{dem} , to the likelihood ratio, LRI_1 , may be expressed as,

$$\begin{aligned}
C_{dem} = & \left(\frac{2}{6}\right) \{LRI(c_{ki} \neq 0; f_{hi} \neq 0; d_{li} \neq 0) - LRI(c_{ki} \neq 0; f_{hi} \neq 0; d_{li} = 0)\} \\
& + \left(\frac{1}{6}\right) \{LRI(c_{ki} \neq 0; f_{hi} = 0; d_{li} \neq 0) - LRI(c_{ki} \neq 0; f_{hi} = 0; d_{li} = 0)\} \\
& + \left(\frac{1}{6}\right) \{LRI(c_{ki} = 0; f_{hi} \neq 0; d_{li} \neq 0) - LRI(c_{ki} = 0; f_{hi} \neq 0; d_{li} = 0)\} \\
& + \left(\frac{2}{6}\right) \{LRI(c_{ki} = 0; f_{hi} = 0; d_{li} \neq 0) - LRI(c_{ki} = 0; f_{hi} = 0; d_{li} = 0)\}
\end{aligned}$$

$$C_{dem} = \left(\frac{2}{6}\right) (LRI_1 - LRI_4) + \left(\frac{1}{6}\right) (LRI_3 - LRI_7) + \left(\frac{1}{6}\right) (LRI_2 - LRI_6) + \left(\frac{2}{6}\right) (LRI_5). \quad (11)$$

It is then easy to verify that,

$$C_{circ} + C_{eff} + C_{dem} = LRI(c_{ki} \neq 0; f_{hi} \neq 0; d_{li} \neq 0). \quad (12)$$

In other terms, by taking into account all possible combinations of the sets of explanatory variables we can then easily derive the respective contributions of circumstances, effort levels and demographic variables to the actual likelihood ratio (see, Shorrocks, 2013).

5. Results of the empirical investigation

This section presents the results of our empirical analysis. Following the approach of Trannoy *et al.* (2010), we first show the results of the two-step procedure in order to identify the determinants of the individual's health status.

This section then presents the results of the Shapley decomposition which allows measuring the contribution of circumstances, efforts and demographic variables to the Pseudo R-square of the health level logit and health ordered logit regressions.

5.1. The regression results

Table 2 gives the results of the five regressions that have been estimated. Columns (2), (3) and (4) give the results of three logit regressions where the dependent variables are respectively the educational level of the individual, whether he/she does not smoke and, whether he/she does have a regular physical activity, the three effort variables. In column (5) we present the results of a logit regression whose dependent variable is the individuals' health status. This variable is equal to 1 when individuals report to have good or very good health and to 0 otherwise. Finally, column (6) presents the results of a health ordered logit regression. At the difference of the logit regression previously mentioned, the dependent variable is the individuals' health status but it is an ordinal variable which can take five possible values: very poor health (1); poor health (2); faire health (3); good health (4); and

very good health (5). The main reason for estimating this last regression is to check the robustness of our model.

As explained previously, the residual of the first logit regression (whose dependent variable is the probability of having a high educational level) was introduced as explanatory variable in the regressions whose dependent variable were the probability that an individual does not smoke, the probability that an individual has a regular physical activity and the health variable (binary and ordinal variable). Similarly, the residual of the logit regression whose dependent variable is the probability that the individual does not smoke was introduced as explanatory variable in the regressions whose dependent variables were the probability that an individual has a regular physical activity and the health variable (binary or ordinal variable). Finally, the residual of the regression whose dependent variable is the probability that an individual has a regular physical activity was introduced as an explanatory variable in the health regressions (with the dependent variable binary or ordinal).

Looking at the second column of Table 2 we observe that males have a higher probability of having a higher education, that the latter decreases with age but at a decreasing rate, is lower among immigrants from Portugal but higher among immigrants from other EU-15 countries and also higher for those individuals whose parents were born in Luxembourg. This probability is also higher when the educational level of the father, as well as that of the mother, is higher. Finally, it is also higher when there were no financial problems in the family when the individual was young.

The third column of Table 2 shows that the probability of not smoking is higher among females, among individuals whose father and mother had a higher education, among those who did not have financial difficulties when they were young and also among those individuals with a higher education level.⁹

In the fourth column we look at the determinants of a regular physical activity and observe that the probability of having such an activity is lower among immigrants from Portugal but higher among immigrants from other EU-15 countries than among individuals born in

⁹ Remember that since the educational level of the individual answering the questionnaire is in this equation measured via the residual of the educational equation, it is orthogonal to the other explanatory variables of the “non-smoking” equation of column 3.

Luxembourg or in non EU-15 countries. It is also lower among individuals who have one parent born in Luxembourg. Finally, we can see that the residuals of the educational and “non- smoking” equations have a positive impact on the probability of having a regular physical activity.

Let us finally look at the two last columns of Table 2 which present the results of the ‘*health*’ equations. Considering health as a binary or an ordinal variable does not make any big difference. In particular, concerning the demographic characteristics, being a man or a woman does not have an impact on the health status of an individual. In contrast, bad health increases with age (but at a decreasing rate).

Looking at the variables reflecting circumstances, we observe that when compared to the health of the natives, the level of health is lower among individuals who immigrated from Portugal but higher among individuals who came from other EU-15 countries. Note also that bad health is less common, the higher the educational level of the father, while the educational level of the mother does not have a significant effect. Additionally, the probability of having good health is higher for those individuals who did not have financial difficulties when they were young.

Lastly, considering the residuals of the effort variables, the three variables have a significant effect: having a higher educational level, not smoking and having a regular physical activity have all a positive impact on health. Remember that the goal of introducing these residuals of the three logit regressions was to have pure measures of effort.

5.2. Results of the Shapley decomposition

Table 3 and Table 4 present the results of the Shapely decomposition. As it was the case in the regressions results, there is no important differences between the results derived on the basis of the health regression where the dependent variable is binary and those obtained when the health variable is an ordinal variable.

In particular, in the case of a binary health variable, circumstances explain 27%, efforts 26% and demographic variables 47% of the pseudo R-square while when the health variable is an ordinal variable, circumstances explain 28%, efforts 20% and demographic variables 52% of the pseudo R-square.

In other words, ignoring the role of demographic variables that are considered biological determinants of an individual's health status, we can conclude that circumstances and efforts have a similar impact on inequality in health in Luxembourg.

6. Concluding comments

Following previous work by Rosas Dias (2009; 2010), Trannoy *et al.* (2010), Jusot *et al.* (2013), Bricard *et al.* (2013) and Lazar (2013), this paper attempted to measure inequality in health opportunity in Luxembourg. More precisely it tried to determine the contribution of circumstances, efforts (and lifestyle) and demographic variables (age and gender) to the overall inequality in health in this country. Health was measured via the answers given to a question on self-assessed health and was considered first as a binary variable, then as an ordinal variable. Several circumstance variables were taken into account: the educational level of each parent, the financial situation of the family during childhood and the area of birth (four areas were considered: Luxembourg, Portugal, other EU-15 countries and other countries). As effort and lifestyle variables we had information on the educational level of the individual, whether he/she smoked and whether he/she had a physical activity on a regular basis. Following Trannoy *et al.* (2010), we adopted an econometric approach which guaranteed that these effort and lifestyle variables were orthogonal to the circumstances variables.

To measure the respective impacts of the three categories of explanatory variables (circumstances, effort and demographic variables) on health inequality we implemented the so-called Shapley decomposition. More precisely, we studied the effect each of these three sets of explanatory variables had on the pseudo R-square of the logit regression whose dependent variable was health (whether a binary or ordinal variable). The results of the empirical investigation indicate that differences in circumstances and effort and lifestyle

explain each around a quarter of the pseudo R-square, the other half of this pseudo R-square being due to differences in demographic variables (gender and age).

Table 1. Summary statistics

<i>Gender</i>	<i>100</i>	<i>Education level mother</i>	<i>100</i>
Men	48	None, primary or secondary	93
Women	52	Post-secondary	7
<i>Age</i>	<i>100</i>	<i>Education level father</i>	<i>100</i>
25-34	20	None, primary or secondary	83
35-44	24	Post-secondary	17
45-54	27	<i>Financial situation</i>	<i>100</i>
55-65	29	Often, most of the time	22
<i>Country of birth</i>	<i>100</i>	Never, occasionally	78
Luxembourg	61	<i>Education level</i>	<i>100</i>
Portugal	11	None, primary or secondary	70
Other EU-15	22	Post-secondary	30
Other	6	<i>Smoke</i>	<i>100</i>
<i>Years of immigration</i>	<i>100</i>	Every day or time to time	24
Natives	61	Does not smoke	76
Less than 10 years	12	<i>Physical activity</i>	<i>100</i>
Between 11 and 24 years	13	Sometimes, never	66
More than 25 years	14	In regular terms	34
<i>Country of birth mother</i>	<i>100</i>	<i>Self-assessment health</i>	<i>100</i>
Luxembourg	44	Very good	29
Portugal	12	Good	47
Other EU-15	27	Fair	17
Other	17	Poor	6
<i>Country of birth father</i>	<i>100</i>	Very poor	1
Luxembourg	46		
Portugal	12		
Other EU-15	25		
Other	17		
Sample size	2332		100

Table 2: Results using logit regressions

Explanatory variables	Dependent Variable: level of education	Dependent variable: does not smoke	Dependent variable: does physical activity	Dependent variable: binary health level	Dependent variable: ordered health
Gender	0.4132 (0.1061)	-0.4598 (0.1009)	-0.0243 (0.0912)	0.1184 (0.1043)	0.0283 (0.0793)
Age	-0.1469 (0.0429)	-0.0486 (0.0401)	-0.0272 (0.0368)	-0.2313 (0.0451)	-0.1767 (0.0320)
Square of age	0.0011 (0.0005)	0.0008 (0.0004)	0.0003 (0.0004)	0.0019 (0.0005)	0.0013 (0.0003)
Country of birth: Portugal	-1.9312 (0.3504)	0.1078 (0.1813)	-1.1417 (0.2194)	-0.7654 (0.1882)	-0.7033 (0.1510)
Country of birth: other EU-15 country	1.3335 (0.1584)	0.0283 (0.1584)	0.5267 (0.1416)	0.4257 (0.1695)	0.4931 (0.1278)
Country of birth: non EU-15 country	0.3669 (0.2388)	0.4548 (0.2602)	-0.0272 (0.2198)	-0.1780 (0.2450)	-0.0933 (0.1910)
Parents born in Luxembourg	-0.2287 (0.1479)	0.0131 (0.1395)	-0.2857 (0.1279)	-0.3498 (0.1450)	-0.2614 (0.1118)
Educational level of father	1.2957 (0.1387)	0.3390 (0.1628)	0.0190 (0.1316)	0.4519 (0.1707)	0.3721 (0.1185)
Educational level of mother	1.8339 (0.2559)	0.7302 (0.2750)	0.2729 (0.1903)	0.1872 (0.2764)	0.1692 (0.1769)
Financial situation	0.3977 (0.1518)	0.4393 (0.1267)	-0.0807 (0.1248)	0.4157 (0.1311)	0.4483 (0.1069)
Constant	2.2971 (0.9075)	1.5170 (0.8548)	0.7944 (0.7944)	7.5628 (1.0345)	
Residual of educational equation		0.3339 (0.0558)	0.2152 (0.0435)	0.2333 (0.0560)	0.3711 (0.0997)
Residual of smoking equation			0.2963 (0.0489)	0.1476 (0.0494)	0.3711 (0.0955)
Residual of physical activity equation				0.3824 (0.0571)	0.3675 (0.0874)
Observations	2,332	2,332	2,332	2,332	2332
Maximum Likelihood	-1112.92	-1222.44	-1413.21	-1143.35	-2655.9
Pseudo R-square	0.219	0.047	0.050	0.114	0.0707

Robust standard errors in parentheses. Significant values in **bold**.

Table 3: Contributions of the circumstances, the efforts and the demographic variables to the Pseudo R-square of the health level logit regression

Contribution	Shapley Contribution of circumstances	Shapley Contribution of efforts	Shapley Contribution of the demographic variables	Pseudo R-square
Absolute Contribution	0.031	0.029	0.054	0.114
Relative Contribution (in %)	27.06	25.56	47.39	100

Table 4: Contributions of the circumstances, the efforts and the demographic variables to the Pseudo R-square of the health level ordered logit regression

Contribution	Shapley Contribution of circumstances	Shapley Contribution of efforts	Shapley Contribution of the demographic variables	Pseudo R-square
Absolute Contribution	0.020	0.014	0.036	0.071
Relative Contribution (in %)	28.08	20.32	51.59	100

References

- Arneson, R.J. (1989), "Equality and equal opportunity of welfare", *Philosophical Studies* 56: 77-93.
- Barry, B. (2005), *Why Social Justice Matters*, Cambridge: Policy Press.
- Berger, F. (2008), "Zoom sur les Primo-Arrivants Portugais et leurs Descendants, " *Vivre au Luxembourg* # 49, CEPS/INSTEAD, Esch-sur-Alzette, Luxembourg.
- Blaxter, M. (1985), *A Comparison of Measures of Inequality in Morbidity*. In John Fox, *Health Inequalities in European Countries*, Aldershot, Gower.
- Bricard, D., Jusot, F., Trannoy, A. and S. Tubeuf (2013), "Inequality of opportunity in health and natural reward: evidence from European countries", *Research on Economic Inequality*, Volume 21: Health and Inequality: 335-370.
- Chantreuil, F. and A. Trannoy (2013), "Inequality Decomposition Values: The Trade-Off Between Marginality and Consistency," *Journal of Economic Inequality*, 11(1): 83-98.
- Checchi, D. and V. Peragine (2010), "Inequality of Opportunity in Italy," *Journal of Economic Inequality* 8: 429-450.
- Cohen, G.A. (1989), "On the Currency of Egalitarian Justice", *Ethics* 99: 906-944.
- DeSalvo, K.B., Fan, V.S., McDonell, M.B. and S.D. Fihn (2005), "Predicting Mortality and Healthcare Utilization with a Single Question," *Health Services Research*, 40(4): 1234-1246.
- Dworkin, R. (1981), "What is equality? Part I: Equality of Welfare", *Philosophy and Public Affairs* 10: 185-246.
- Dworkin, R. (2000), *Sovereign Virtue: The Theory and Practice of Equality*, Cambridge: Harvard University Press.
- European Commission (2011), "Migrants in Europe – A Statistical Portrait of the First and Second Generation", Statistical Book, Eurostat, The Statistical Office of the European Union, Luxembourg.
- Fleurbaey, M. (2008), *Fairness, Responsibility and Welfare*, Oxford University Press: Oxford.

- Fleurbaey, M. and E. Schokkaert (2009), “Unfair Inequalities in Health and Health Care,” *Journal of Health Economics* 28: 73–90.
- Greene, W. H. (1992) LIMDEP Version 6.0: User’s Manual and Reference Guide, Econometric Software Inc., New York.
- Idler, E.L. and Y. Benyamini (1997), “Self-related Health and Mortality: A Review of Twenty Seven Community Studies,” *Journal of Health and Social Behaviour*, 38: 21-37.
- Jusot, F., Tubeuf, S. and A. Trannoy (2013), “Circumstances and Efforts: How Important is their Correlation for the Measurement of Inequality of Opportunity in Health?,” *Health Economics* 22: 1470-1495.
- Lazar, A. (2013), “Ex-ante and Ex-post Measurement of Inequality of Opportunity in Health: Evidence from Israel,” *Research on Economic Inequality*, vol. 21, *Health and Inequality*, P. Rosa Dias and O. O’Donnell editors, pp. 371-395.
- Lefranc, A., N. Pistolesi and A. Trannoy (2006), “Equality of opportunity: definitions and testable conditions, with an application to income in France,” ECINEQ (Society for the Study of Economic Inequality) Working Paper 2006-53.
- Lefranc, A., Pistolesi, N. and A. Trannoy (2009), “Equality of Opportunity and Luck: Definitions and Testable Conditions, with an Application to Income in France,” *Journal of Public Economics* 93 (11-12): 1189-1207.
- Roemer, J. E. (1995), “Equality and Responsibility” in *Boston Review: A Political and Literary Forum*. Available at <http://bostonreview.net/BR20.2/roemer.html>
- Roemer, J. E. (1998), *Equality of Opportunity*. Cambridge: Harvard University Press.
- Rosa Dias, P. (2009), “Inequality of Opportunity in Health: Evidence from a UK Cohort Study,” *Health Economics* 18(9): 1057–1074.
- Rosa Dias, P. (2010), “Modelling Opportunity in Health under Partial Observability of Circumstances,” *Health Economics* 19(3): 252–264.
- Sastre, M. and Trannoy, A. (2002) “Shapley inequality decomposition by factor components: some methodological issues,” *Journal of Economics* 9, 51–90.
- Sermet, C. and E. Cambois (2002), *Mesurer l’Etat de Santé*. In Caselli G., Vallin J., Wunsch G., Démographie : analyse et synthèse. Les déterminants de la mortalité, Paris, Editions de l’INED, 2002/12, Vol. III, Chapitre 41, pp. 25-52.

- Shorrocks, A.F. (2013), "Decomposition Procedures for Distributional Analysis: A Unified Framework Based on the Shapley Value," *Journal of Economic Inequality*, 11, 1, pp. 99-126.
- Swift, A. (2005), "Justice, luck and the family: normative aspects of the intergenerational transmission of economic status", In Bowles S., Gintis H., Osborne-Groves M. (ed.), *Unequal Chances: Family Background and Economic Success*. Princeton University Press: Princeton: 256-276.
- Trannoy, A., Tubeuf, S., Jusot, F. and M. Devaux (2010), "Inequality of Opportunities in Health in France: A First Pass," *Health Economics*, 19(8), 921-938.
- Tubeuf, S., Jusot, F., Devaux, M. and C. Sermet (2008), "Social Heterogeneity in Self-reported Health Status in Measurement of Inequalities in Health", *Working Paper IRDES Series* 12.

