Palacký University Olomouc

University of Clermont Auvergne

University of Pavia

MASTER THESIS

Inequality of Opportunity in Adult Health: Empirical Evidence from Mexico

Yanelkis Fernández-Molina

Supervisor: Cinzia Di Novi

GLODEP 2020

Inequality of Opportunity in Adult Health: Empirical Evidence from Mexico

Yanelkis Fernández-Molina

Supervisor: Cinzia Di Novi

GLODEP 2020

Declaration

I, hereby, declare that the Thesis ‘Inequality of Opportunity in Adult Health: Empirical Evidence from Mexico’, submitted to the GLODEP Consortium 2020, is my original work, and any theoretical and empirical literature, as well as the dataset used for the analysis, have been cited and referenced.

Yanelkis Fernández-Molina

Date: June 7th, 2020

UNIVERZITA PALACKÉHO V OLOMOUCI

Přírodovědecká fakulta

Akademický rok: 2019/2020

ZADÁNÍ DIPLOMOVÉ PRÁCE

(projektu, uměleckého díla, uměleckého výkonu)

Jméno a příjmení:	**Yanelkis FERNANDEZ**
Osobní číslo:	**R180544**
Studijní program:	**N1301 Geography**
Studijní obor:	**International Development Studies**
Téma práce:	**Inequality of Opportunity in Health: Empirical Evidence from Mexico**
Zadávající katedra:	**Katedra rozvojových a environmentálních studií**

Zásady pro vypracování

Background

In the last decades, Mexico has become one of the most important economy in the Latin American and Caribbean region. Its economic development has been remarkable, transferring part of this progress to its population in terms of access to basic goods and services. However, Mexico continues to be one of the most inequal countries in that region, which is mainly reflected in the differences in access and outcomes on people's welfare. As an illustration, the Human Development Index (HDI) of Mexico for 2018 was about 0.767, positioning it at 76 out of 189 countries; however, when inequality is discounted from the value , it losses 22.5% of the value.[1] When zooming to the dimensions of the HDI, the loss due to inequality on the health dimension of the index , assessed by the life expectancy, is of 10.5%, representing a phenomenon that needs to be studied in more depth (United Nation Development Programme, 2019).

Aims of the Thesis

According to Roemer's framework, the achievements of people in terms of welfare, in this case health, depends on two aspects: „circumstances" and „effort". The *circumstances* refer to the characteristics of the individual that are not under its responsibility (e.g., sex, gender, parental socioeconomic status, race) and are sources of questionable social inequalities. On the other hand, *efforts* refer to the characteristics of the individuals that they are responsible for, and depend on their behaviour and decisions (e.g., choosing a proper diet, not smoking, drinking or consuming drugs, etc.) (Rosa Dias, 2010; Carrieri & Jones, 2018; Gallardo et. al, 2017).

In order to study in more depth these elements, the thesis aims to assess the existence of inequality of opportunity in the health status of the Mexican population. Specifically, our objective is to analyse the presence of health inequalities related to the socioeconomic and sociodemographic dimensions that individuals cannot control, taking into account the effort an individual exerts in order to have a healthy life.

Data and Methodology

In order to identify the existence of inequality of opportunity, we need to decompose the part of the health outcome that is attributable to circumstances and the part outcome that is attributable to efforts. We will apply a regression-based decomposition method of health inequality in order to assess the contribution of each variables chosen as indicator of circumstances and to proxy effort in the health outcomes of Mexicans. We will use data from the National Health and Nutrition Survey of Mexico in its latest version, 2018. This survey collects information on more than 29,000 individuals at the national level and represents a solid source of information to obtain representative information about the population in Mexico.

[1] HDI was developed by the United Nations to measure countries' levels of social and economic development . It includes three principal areas of interest of human development: living standards (Gross National Income per capita), population educational level (mean and expected years of schooling), and health (life expectancy at birth).

Rozsah pracovní zprávy:	
Rozsah grafických prací:	
Forma zpracování diplomové práce:	**tištěná**
Jazyk zpracování:	**Angličtina**

Seznam doporučené literatury:

Carrieri, V., & Jones, A. (2018). Inequality of opportunity in health: A decomposition-based approach. *Wiley*.
Gallardo, K., Varas, L., & Gallardo, M. (2017). Inequality of opportunity in health: evidence from Chile. *Revista de saude publica*, 51 – 110.
Jones, A., Rice, N., & Rosa Dias, P. (2012). Quality of schooling and inequality of opportunity in health. *Springer*, 369?394.
O'Donnell, O., van Doorslaer, E., Wagstaff, A., & Lindelow, M. (2008). *Analyzing health equity using household survey data: a guide to techniques and their implementation.* Washington: World Bank.
Rosa Dias, P. (2007). Giving equality of opportunity a fair innings. *Wiley InterScience*, 109 – 112.
Rosa Dias, P. (2010). Modelling opportunity in health under partial observability of circumstances. *Wiley Interscience*, 252-264.
Sanoussi, Y. (2017). Measurement and analysis of inequality of opportunity in access of maternal and child health care in Togo. *Health Services Research*.
UNDP (2019). Human Development Report 2019: Inequalities in Human Development in the 21st Century. Country profiles: Mexico. Retrieved from: http://hdr.undp.org/sites/all/themes/hdr_theme/country-notes/MEX.pdf

Vedoucí diplomové práce: **prof. Cinzia Di Novi**
Katedra rozvojových a environmentálních studií

Datum zadání diplomové práce: **13. ledna 2020**
Termín odevzdání diplomové práce: **22. května 2020**

L.S.

doc. RNDr. Martin Kubala, Ph.D.
děkan

doc. RNDr. Pavel Nováček, CSc.
vedoucí katedry

V Olomouci dne 28. ledna 2020

Acknowledgments

I would like to thank the guidance and support that I received from Professor Cinzia di Novi during this process. I really appreciate the unquantifiable effort that she dedicated to review, comment, and offer solutions for this Thesis. Additionally, her lectures on Health Economics were the main motivation for deciding my research topic. During classes, she transmitted her rich knowledge on the area, and unconsciously encourage me to study inequalities and inequities in health. Without Prof. di Novi, I would not have explored this captivating world of health economics.

I must thank the GLODEP Consortium for giving me the opportunity to enroll in this program. It has been a growing experience in academic and personal terms. The quality of the lectures and lecturers is unquestionable, which have allowed me to expand my theoretical and practical knowledge on Development Studies.

I am thankful for being part of the second cohort of GLODEP, specially because of my colleagues. They are tremendous professionals who shared humbly their knowledge and experiences throughout these two years. However, their best attribute is that they are great friends who never doubted on offering their support in times of need.

Last but most important, I would like to thank my mother and father for all their effort to overcome circumstances and giving me more than what they could. I thank them for being always there for me, and for supporting the choices that I made that brought me to this program.

"Por último, pero más importante, quisiera agradecer a mi madre y a mi padre por todo el esfuerzo que ellos pusieron para superar circunstancias adversas y brindarme más de lo que ellos podían dar. Les agradezco por siempre estar ahí para mí, y por apoyarme en las decisiones que me condujeron a este programa."

Thank you,

Yanelkis Fernández-Molina

Abstract

The prevalence and mortality of non-communicable diseases (NCD) have become the most challenging health problem in Mexico, with significant differences arising from gender, age, ethnic groups, regions, zone of residence, and others. According to the Equality of Opportunity (EOp) paradigm, inequality arising from these aspects are unfair and societies should compensate for them. This Thesis analyzes the inequality in adult health due to unfair circumstances by applying a regression-based approach and the Mean Logarithmic Deviation to data from Mexican National Health and Nutrition Survey 2018. Adult health status is measured by the glycated hemoglobin (HbA1c), albumin in the blood, and the systolic blood pressure, which are indicators used to diagnose and monitor the deadliest NCD's, i.e. cardiovascular and liver diseases, and diabetes.

The results evidenced significant *lower-bound* levels of unfair inequality of almost a quarter for SBP, 16.6% for HbA1c, and 13.4% for albumin in the blood. The Shapley decomposition revealed that sex, age, and hereditary conditions are behind most part of the inequality of opportunity (IOp). The socioeconomic conditions and the region in which the individuals are born, shape the levels of IOp as well. Finally, the estimations for males and females suggest that not only differences between them are associated to large shares of unjust inequality in health, but also the opportunities they face within themselves are unequal. In consequence, in order to 'level the playing field' the Mexican society should compensate individuals for the aforementioned aspects, because these inequalities are unfair, thus, people should not be held responsible for them.

Keywords: Equality of Opportunity, health inequality, biomarkers, non-communicable diseases, mean logarithmic deviation, Shapley decomposition

Contents

List of Graphs

List of Tables

List of Appendices

List of Abbreviations

CVD	Cardiovascular diseases
ENSANUT	National Health and Nutrition Survey of Mexico
EOp	Equality of Opportunity
HbA1c	Glycated hemoglobin
HDI	Human Development Index
INEGI	National Institute of Statistics and Geography
INSP	National Institute of Public Health
IOp	Inequality of Opportunity
MLD	Mean Logarithmic Deviation
NCD	Non-communicable diseases
PND	National Development Plan of Mexico
SBP	Systolic blood pressure
UNDP	United Nations Development Program
WHO	World Health Organization

Introduction

Health is undoubtedly an important dimension of welfare, but its direct impact on the other spheres of wellbeing makes it even more relevant. Bad health is highly detrimental for people, and its unequal distribution usually translates to inequalities in other dimensions, with the possibility of exacerbating them (World Bank, 2006). While the negative consequences of bad health and its unequal distribution are undisputed, the roots of this inequality and where it stems from have been widely discussed.

Some theorists of equity and social justice consider as unfair the differences in outcomes arising from elements people cannot control, such as age, gender, ethnicity, place of birth, socioeconomic status of the family they were born into, and others. These aspects are called *circumstances*. In contrast, differences arising from conscious decisions and actions are fair because people have control over them. These decisions are called 'efforts' (Roemer, 1998).

This branch of thought is called the Theory of Equality of Opportunity (from now, EOp). The EOp is based on the compensation principle, which states that "*inequality due to circumstances beyond individual responsibility is ethically unjustified*" (Ferreira & Peragine, 2015, p. 9). This kind of inequality is called 'inequality of opportunity (IOp)', and individuals should be compensated for it because it is unfair (Fleurbaey, 1994). The compensation principle is divided into two approaches: the ex-ante and ex-post compensation. The ex-ante conceives compensation from unfair aspects before people exert effort, and the equality condition is that everyone has an equalized set of opportunities. While the ex-post compensation states that everyone who exerted the same level of *effort* should achieve the same outcome (Ferreira & Peragine, 2015).

Taking these aspects into consideration, the United Mexican States becomes an interesting case study to apply the EOp paradigm because of three main reasons. First, despite being placed in the High Human Development category, and having increased its Human Development Index (HDI) in 8.8% since 2000 (UNDP, 2019), high levels of poverty and inequality still prevail with significant gaps between regions (OECD, 2019). Additionally, evidence has shown that Mexico is one of the most unfairly unequal countries in Latin America and the Caribbean -in terms of income-. These inequalities are persistent over the time, are transferred by generations, and arise from differences in gender, ethnic groups, regions, and others (Wendelspeiss-Chávez-Juárez, 2014; Vélez-Grajales, et al.,2018; Monroy-Gómez-Franco et al., 2018).

Second, the prevalence and mortality of non-communicable diseases (NCD) represents one of the biggest health problems in Mexico because of their magnitude, frequency, and their appearance earlier in life (Villalpando & Rodrigo, 2010; Córdova-Villalobos et al., 2008; CONEVAL, 2018). NCD's mortality rate accounts for most of the deaths in the country. Specifically, the main causes of deaths in Mexico are cardiovascular diseases (CVD), diabetes mellitus, and liver diseases, representing 41.0% of total mortality. Moreover, their growth rates have increased constantly over the past two decades (CONEVAL, 2018).

Third, studies have evidenced that the socioeconomic background in which individuals are born influences the abilities and opportunities they count with to achieve better outcomes, and that the health conditions of parents influence the health endowment of individuals (Campos-Vázquez, 2015; Moreno-Jaimes, 2017).

However, to the best of our knowledge, the theory of EOp has not been applied empirically to measure unfair inequality in *adult health* in Mexico. This gap represents a major hollow because of the reasons stated above. In other words, these aspects translate into the fact that it is unknow how unfairly unequal is the health status of people in a country characterized by unjust income inequality, with increasing sanitary problems in the population, and with intergenerational transmission of inequalities.

In consequence, the following question arises: How unequal is the distribution of adult health status in Mexico due to individual's circumstances?

Estimating the contribution of circumstances to inequality in the indicators -biomarkers- that diagnose and monitor CVD's, diabetes mellitus and liver diseases, becomes imminent. These biomarkers are the systolic blood pressure (SBP), glycated hemoglobin (HbA1c) and albumin in the blood (Campos-Donato et al. 2019; American Diabetes Association, 2019; Peters, 1996).

Therefore, in more specific terms:

1. What is the extent of inequality of opportunity in the distribution of SBP, albumin in the blood, and HbA1c in the Mexican population?
2. What is the contribution of each observed circumstance to inequality of opportunity in the biomarkers chosen?
3. How do circumstances affect separately females and males in distribution of SBP, albumin in the blood, and HbA1c?
4. What are the differences in the partial contributions of each circumstance to the unfair inequality for males and females?
5. What are the circumstances that contribute the most to IOp?

To answer these questions, this Thesis aims to analyze the level inequality of opportunity in adult health status in Mexico by applying the ex-ante approach to Roemer's model on the theory of Equality of Opportunity.

Using the last wave of the National Health and Nutrition Survey 2018, this research has the objective to measure the extent of inequality related to aspects that individuals cannot be held responsible for in the distribution of HbA1c, albumin in the blood, and the systolic blood pressure. Additionally, this study calculates the contribution of each circumstance to the unfair inequality for the overall population of adults, and for males and females separately.

In order to measure the level of IOp in each biomarker, a regression-based approach is applied to estimate the expected health outcomes. Then, using the Mean Logarithmic Deviation (MLD), the point estimates and relative values of IOp are obtained as a residual of the total inequality and the inequality from a counterfactual distribution. The contribution of each circumstance to IOp is obtained by applying a Shapley decomposition.

The rest of the thesis is organized as follows. Chapter I presents the review of previous literature on the theory of Equality of Opportunity along with its empirical application in Mexico as well as in adult health in other countries. Chapter II explains the theoretical framework employed, followed by specifications of the empirical model. Then, Chapter III discusses the description and analysis of results, while the Conclusion summarizes and concludes the thesis with with some final remarks and recommendations.

Chapter I

Literature review

This chapter presents the theoretical and empirical literature relevant to the analysis of inequality of opportunity in health status. First, the Section 1.1 introduces the grounds in which the Equality of Opportunity theory is based and its main approaches and concepts. Second, Section 1.2 presents the studies that have analysed inequality of opportunity in Mexico. Then, Section 1.3 offers a review on the empirical literature specific to in IOp in adult health following an ex-ante compensation perspective. In that section, the databases and the variables chosen, along with the methodological approaches are explained. Finally, Section 1.4 discusses the existing literature gap and how this study contributes to fill it.

1.1 Review of theoretical literature

1.1.1 The Theory of Equality of Opportunity

The analysis of equity in societies has evolved significantly over the decades. Traditionally, under the 'welfarist approach', social justice was conceived as the equal distribution of either utility, welfare or preference satisfaction. Many critics of this paradigm challenged the idea of using the distribution of the aforementioned aspects as the unique way to measure social justice (Ferreira & Gignoux 2011; Ferreira & Peragine, 2015). Additionally, they argued that this paradigm ignores individual's responsibility on their choices, characterizing this kind of equality as *non-ethically* desirable by society (Lefranc et al. 2006; Ferreira & Peragine, 2011; Roemer & Trannoy , 2013).

For these critics, societies should focus on a different *equalisandum* such as 'primary goods' proposed by John Rawls, 'capabilities and functionings' by Amartya Sen, 'resources' by Ronald Dworkin, 'opportunity to welfare' by Richard Arneson, and 'access to advantage' by Gerald Cohen (Fleurbaey, 1994; Roemer, 1996 & 2002; Roemer & Trannoy , 2013; Ferreira & Peragine, 2015). [1] These contributions moved the focus from *equality of outcomes* to *equality of opportunities* (Roemer, 1998; Lefranc et al., 2006; Rosa Dias, 2009). Moreover, they brought to the discussion the idea of people's responsibilities as an aspect to consider in order to have desirable and ethical equity in societies (Roemer & Trannoy , 2013). This set of contributions is called Theory of Equality of Opportunity (from here EOp).

[1] In words of Cohen (2011, p. 5) "an equalisandum claim specifies that which ought to be equalized, what, that is, people should be rendered equal in".

The EOp theory was later divided into two families of approaches. The first one, called "direct approach", implies equalizing a set of opportunities to all individuals. In other words: "*rendering the sets of choices available to different individuals the same*" (Roemer, 2002). It tries to measure the extent of inequality in a society by ranking individuals according to the set of opportunities they are endowed with. The direct approach implies that inequality is reduced if the distribution of the opportunity basket along the population is improved. It does not put much focus on individual's responsibilities or the outcomes (Pignataro, 2012).[2]

Roemer (2002) argues that this approach has failed to be put into practice in terms of empirical works referring to how it should be applied in real-life. Nevertheless, the direct approach opened the arena for the indirect one, which was mainly influenced by Roemer's contributions and the metaphor "levelling the playing field" (Roemer, 2002; Ferreira & Gignoux, 2015)

1.1.1.a The indirect approach to equality of opportunity

The indirect approach focuses on the *outcomes* of "*a given distribution of opportunities as manifested in the joint distribution of an observable advantage and a number of individual characteristics*" (Ferreira & Peragine, 2015, p. 8). One of the main discussions in the indirect approach to EOp is what part of inequality is legitimate or not, fair or unfair, just or unjust (Jusot et al., 2013). Roemer (1998) states that societies should focus on 'levelling the playing field', i.e. giving people the resources they require, so everyone has the same opportunities. The aspects to 'level' are the "*differential circumstances of individuals for which we believe they should not be held accountable, and which affect their ability to achieve or have access to the kind of advantage that is being sought*." (Roemer, 1998, p. 5).

In order to put this into practice, the model asserts that the outcome of individuals should be divided into the share of it due to *circumstances* and due to *efforts* (Roemer, 1998, 2002, 2016). 'Circumstances' refer to elements that individuals were born into, do not control, and are not responsible for, such as gender, race, place of birth, socioeconomic status of the family they were born into, the quality and quantity of services in the environment, etc. (Roemer, 1998) . The inequality arising from these are unfair. However, the differences in outcomes due to voluntary, conscious choices, and to actions controlled by individuals should not be subject for compensation because they are fair. These aspects are called 'efforts'. However, these 'voluntary choices' can be also influenced by circumstances. (Roemer, 1998; Pignataro, 2012; Ferreira & Peragine, 2015)

A major concern arises in the application of the EOp analysis. Ferreira and Peragine (2015) state that the EOp theory is an "empty box" without the definition of which aspects are under individual responsibility and not. Even though, Roemer highlighted some circumstances, he exerts that "*In actual practice, the society in question*

[2] Ferreira and Peragine (2015) argue that this set of opportunities is similar to the array of capabilities that Sen identifies in his work, "from which individuals choose a particular vector of functionings."

shall decide, through some political process, what it wishes to deem 'circumstances'" (Roemer, 1998, p. 8). The data availability will be a key determinant for these choices as well (Roemer, 1998; Ferreira & Peragine, 2015).

After identifying which aspects should be considered as circumstances and as efforts, Roemer proposes to divide the population into groups of individuals that share the same set of circumstances. These groups are called 'types' or 'social types'. All individuals that belong to the same type, theoretically, share equal capacity to achieve a certain outcome, since individuals within a type share the same circumstances. *Inequality of opportunity* (from here IOp) exists when there are differences *between* these groups, because under equality, circumstances should not matter for the outcome, only the effort exerted by people (Roemer, 1998). For a more detailed explanation of the theoretical model of the indirect approach to EOp, see Appendix 1.

The principle of compensation

One of the principles that the EOp theory sets its ground is the compensation principle. It states that "*inequality due to circumstances beyond individual responsibility is ethically unjustified*" (Ferreira & Peragine, 2015, p. 9). It implies the compensation to individuals because of differences in their achievements arising from aspects that were not under their control (Fleurbaey, 1995).

Within the compensation principle, the literature proposes two approaches: the ex-ante and ex-post compensation. The ex-ante focuses on the differences in outcomes arising from the different opportunities people faced, evaluating their *circumstances*. The ex-post focuses on evaluating the differences in outcomes in people that exerted the same level of *effort*. (Fleurbaey & Peragine, 2013; Ramos et al. 2015; Ferreira & Peragine, 2015).

There is no agreement in the literature on which approach is better (Fleurbaey & Peragine, 2013). However, the difficulties in observing efforts due to the lack of data in most surveys, has influenced the empirical applications of the *ex-post* approach (Wendelspiess-Chávez-Juárez, 2014). The most found estimations of IOp are from an *ex-ante* approach, suggesting that there is more consensus on measuring unfair inequality observing circumstances (Wendelspiess-Chávez-Juárez, 2014; Ferreira & Peragine, 2015).

This thesis follows an ex-ante compensation approach to measure IOp in health status in Mexico. The ex-ante approach was chosen because of the larger availability of methods in the literature, and because there is more consensus for its implementation (Wendelspeiss & Soloaga, 2014).

1.2 Review of the Empirical Literature

1.2.1 Empirical literature on the ex-ante compensation approach to Inequality of Opportunity in Mexico

The EOp theory has been applied to areas such as education, income, welfare, unemployment, etc. In the past years, there has been increasing interest from international organizations, policy makers and researchers to analyze IOp in health (Rosa Dias, 2009 & 2010; Jusot et al., 2013; Fajardo, 2016). However, to the best of our knowledge, there is no study that measures IOp in adult health in Mexico. The empirical literature in the country has focused on measuring the extent on unfair inequalities in income, wealth, and education.

In terms of dependent variables, the household assets index has been the outcome mostly analyzed. However, Wendelspeiss-Chávez-Juárez (2014) computes the IOp analysis for eight different dependent variables to proxy income and wealth to evidence the impact in the results of the variable chosen. Earlier works aimed to analyze IOp in education as well (Paes de Barro et al. 2009; Soloaga & Wendelspiess, 2010).

The set of circumstances is very homogeneous for the most empirical works. They usually include parent's education, sex, the individual's socioeconomic status in the childhood, zone of residence (urban and rural), indigenous condition of the person, the current wealth index, and others. The exception is Soloaga & Wendelspiess (2010) which included the education of the household head, access to public services and assets ownership of the households, and the number of siblings of the individual to measure the level of IOp in educational attainment and child labor.

All studies implement *regression-based approaches* to avoid the lack of observations (and its statistical consequences) that brings dividing the population into '*types'*. The difference in terms of methodology is that Paes de Barros et al. (2009) and Soloaga & Wendelspiess (2010, 2013) use probit models to compute how circumstances relate to the outcome chosen, then the dissimilarity index or the mean logarithmic deviation (MLD) are applied to measure the levels of inequality. In contrast, the other group of studies implement Ordinary Least Squares (OLS) for continuous variables, and the MLD to obtain *lower-bound* point estimates of IOp (Wendelspeiss-Chávez-Juárez, 2014; Vélez-Grajales, et al.,2018; Monroy-Gómez-Franco et al.,2018; Monroy-Gómez-Franco & Corak, 2019).

The results obtained are homogeneous across studies: the wealth of the household of origin, parent's education and occupation, zone of residence (urban and rural), and indigenous condition, account for the largest shares of IOp. In contrast, sex is associated to low shares of IOp. In general, the level of unfair inequality is estimated around 40% of the outcome of interest, with some variations when authors explore the inclusion of other circumstances to the analysis. (Wendelspeiss-Chávez-Juárez, 2014; Vélez-Grajales, et al., 2018; Monroy-Gómez-Franco et al., 2018; Monroy-Gómez-Franco & Corak, 2019).

The exception is Soloaga & Wendelspiess (2010), in which the set of circumstances differ significantly from other. For this research, the authors were more concerned on the characteristics of the household of origin. They found that the largest share of inequality of opportunity were due to differences in the education of the father and the access to services of the household.

See *Appendix 2* for a summary on the studies identified that have aimed at analyzing IOp in Mexico.

1.2.2 Empirical applications of the Equality of Opportunity theory in adult health status

The empirical literature of IOp in health mostly focuses in European countries (Trannoy et al., 2009; Rosa-Dias, 2009; Rosa-Dias, 2010; Carrieri & Jones, 2018; Davillas & Jones, 2020; Carrieri et al., 2020). However, in developing countries there have been studies in Indonesia, Colombia, and Chile, to the best of our knowledge (Jusot et al, 2014; Fajardo, 2016; Gallardo et al., 2017).

In general, these studies aimed at measuring the extent of IOp in adult health, by calculating the contribution of circumstances (and efforts in some cases) to inequality in health status. Roemer's model is used as the theoretical framework, then following the ex-ante compensation view. The application of the IOp analysis brings to light significant and persistent inequalities in health due to circumstances in all its empirical applications, but with variations in the methodology applied (Trannoy et al., 2009; Rosa-Dias, 2009; Rosa-Dias, 2010; Jusot et al, 2014; Fajardo, 2016; Gallardo et al., 2017; Carrieri & Jones, 2018; Davillas & Jones, 2020; Carrieri et al., 2020).

For a summarized review, see *Appendix 3,* which displays the main aspects of these studies.

1.2.2.a Source of the data

The data used for these studies correspond mostly to longitudinal surveys, but some cross-sectional surveys are also analyzed. In the cases where panel data was used, there was a wider availability of variables to construct a solid background set on aspects related to the childhood of the cohort members, specially related to socioeconomic circumstances of the parents, as well as health endowment of kids (Rosa-Dias, 2009;Trannoy, et al., 2009; Rosa-Dias, 2010; Jusot et al., (2014).

1.2.2.b Measures of health

The most widely used measure of health in the population is the Self-Assessed Health (SAH). The following standard self-assessed health status question was asked: "Would you say that in general your health is excellent, very good, good, fair, poor?". SAH is supported by a literature that shows the strong predictive relationship between people's self-rating of their health and mortality or morbidity (Idler & Benyamini, 1997; Kennedy et al., 1998). Moreover, SAH correlates strongly with more complex health indices such as functional ability or indicators derived from health service use (Unden & Elofosson, 2006). Gallardo et al. (2017), Fajardo (2016) and Rosa-Dias (2010 & 2009) use SAH's measures as their health outcome variable. Other studies use the existence of a long-term illness, disability or the presence of some mental illness to complement the SAH indicator (Rosa Dias, 2010).

Jusot et el. (2014), asserts that even though it has been demonstrated that this variable is a good predictor of adult health, some limitations have been identified by the literature. First, the reporting bias may arise if in the population, some sub-groups use different categories to evaluate their status of health, although they share an equal level of objective health. Second, as this is an ordinal variable, health cannot be measure in a continuous way, bringing the 'ordinal scale problem' (Jusot et al., 2014).

In recent years, the use of biomarkers was introduced in the analysis of IOp as an objective measure of health (Jusot et al., 2014; Carrieri & Jones, 2018; Davillas & Jones, 2020; Carrieri et al. 2020).[3] These biomarkers are obtained by nurses or specialized interviewers that take blood samples, measure height and weight, and other data. As stated in Carrieri & Jones (2018), the great advantage of using them as indicators of health is that they do not have reporting heterogeneity bias. Additionally, Davillas & Jones (2020) assert that as these are continuous measures and the value of these indicators mean the severity of the condition, the health outcome of interest does not need to be "cardinalized", as what happens with the SRH and SAH measures.

Following this line, Carrieri and Jones (2018) use blood-tests results for the level of cholesterol, fibrinogen, and glycated hemoglobin as biomarkers to measure health status. Additional to the measures of cholesterol and fibrinogen, Davillas & Jones (2020) use other blood tests, the waist-to-height ratio, systolic blood pressure and the allostatic load (i.e. a composite measure containing the sum of all the blood-based biomarkers). Carrieri et al. (2020) uses the allostatic load to estimate IOp.

1.2.2.c Circumstances

The choice of circumstances for the IOp analysis in monetary outcomes (income, earnings, and consumption), has set the grounds for other areas of analysis. Ferreira & Peragine (2015) compared eight (8) studies on the ex-ante IOp approach applied across 41 countries. In 7 out of 8 of the studies, the variables related to parental characteristics, either their education or occupation, or both, as well as region of birth or geographical location are always included, as well as gender, ethnicity, nationality, mother tongue, and even religion.

In the IOp literature in health, the choice of circumstances is similar to the ones aforementioned. As circumstances are aspects that a person cannot control, the literature includes socioeconomic characteristics of the

[3] A biomarker is "a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention" (Biomarkers Definitions Working Group, 2001). They are tested widely to evaluate their influence and association to the diseases they diagnose and monitor.

parents proxied by both 'social class'[4] and educational attainment of the parents of the individual. Additionally, other aspects of the lives of parents are considered, especially the ones related to their risk behaviors as in Rosa Dias (2009 & 2010) which included parental smoking conditions.

Health endowment variables such as birthweight, morbidity and obesity when childhood are included. Moreover, the occurrence of diseases in parents and siblings of the individual, or the longevity of parents, are analyzed to account for hereditary genetic conditions (Rosa Dias, 2009; Trannoy et al., 2009; Rosa Dias, 2010; Jusot et al., 2014).

Sex and age are always considered as circumstances. However, as they are variables that account for a significant and high share of inequality in health, sometimes authors include them as controls (Jusot et al., 2014; Fajardo, 2016), or the estimations are also computed separately (Davillas & Jones, 2020).

For the analyses on developing countries, geographical variables are always included. The zone of residence (urban and rural), and province and region of birth are included in the studies on Indonesia, Colombia and Chile (Jusot et al., 2014; Fajardo, 2016; Gallardo et al., 2017). Fajardo (2016) explains that this is especially important for Colombia because of the differentiated locations where ethnic minorities live.

To account for spatial differentiation, Carrieri and Jones (2018) included the level of deprivation of the neighborhood where individuals currently live in because to modify it, individuals will have to undergo elevated costs in financial, social and psychological ways, therefore it is hard to change.

Ethnicity is another dimension people cannot influence; therefore, it is also considered as a circumstance (Jusot et al. 2014; Fajardo G., 2016).[5] In the case of Davillas & Jones (2020), ethnicity is proxied by a variable that reflects whether the individual spoke English in her childhood at home.

The level of education attained is considered beyond individual's responsibility under the justification that it is highly influenced by parents (Trannoy et al., 2009; Rosa Dias, 2010; Jusot et al., 2014; Fajardo G., 2016; Carrieri & Jones, 2018; Davillas & Jones, 2020). In the specific case of Carrieri & Jones (2018) and Davillas & Jones (2020), the level of secondary schooling achieved by age 18 is considered as a circumstance.

1.2.2.d Efforts

In the health literature, lifestyles aspects are considered to represent the effort exerted by people: exercise, smoking and drinking behaviors, consumption of 'healthy or unhealthy' food, etc. (Jusot et al. 2014; Carrieri &

[4] This is usually proxied by the occupation of the parents in the childhood and adolescence of the individual. See Table 1 for more detailed information on the variables used.

[5] In the case of Jusot et al. (2014), the language spoken in daily life at home is included, while an ethnic group variable in Fajardo (2016).

Jones, 2020). According to Carrieri & Jones (2018), "lifestyles are determined by individual decisions to invest in health capital, and, therefore, they are, at least partly, within individual control".

The variables used to approximate the efforts exerted by people are: Cigarrette smoking, consumption of alcohol, and proxies for (un)healthy diets (Rosa Dias, 2009 & 2010; Carrieri and Jones, 2018). In the case of Rosa Dias (2009), consumption of fried food was the proxy for an unhealthy diet, while for Carrieri and Jones (2018) the consumption of fruit and vegetables was used. Unlike the literature that only relies on self-reported variables for effort, this last paper used the biomarker of saliva cotinine to measure cigarette smoking.

1.2.2.c Methodology implemented

When analysing the different techniques in the empirical literature, three common aspects are found: first, several studies use of the stochastic dominance tests to assess the hypothesis of equality of opportunity in health status. Second, most of them use parametric approaches (i.e. regression-based estimations) to measure the extent of IOp. Third, most treat efforts as unobserved.

First, stochastic dominance tests are used to assess the existence of IOp in the distribution health status. It basically consists in comparing two cumulative distribution functions (CDF), under the hypothesis that one CDF has a better distribution than the other. Rosa Dias (2009), Trannoy et al. (2009), Jusot et al. (2014) and Fajardo (2016) apply a first order stocastic dominance test, while Gallardo et al. (2017) apply a second order one. The Kolmogorov–Smirnov (KS) test is used to assess for first order stochastic dominance (FOSD) between types.

Second, to *quantify* the extent of the IOp, the literature implements a variety of methods. Rosa Dias (2009) implemented two approaches to measure IOp: the Gini-opportunity index[6], which is calculated using 3 social types (one per category of father's occupational category in 1974), and the 'conditional equality', which treats each individual as a type, avoiding "subjectivity" in the definition of the social types defined by Roemer. As the number of types will be the same as the number of individuals, the Gini-opportunity index equals, by computation, the conventional Gini coefficient.

In the case of Trannoy et al. (2009), a multivariate regression[7] is applied to determine the way each circumstance affects overall inequality, referred to as 'channels of transmissions'. Then, the Gini Index is calculated. Rosa Dias (2010) integrates the Grossman's model on health capital and the demand for health to

[6] Gini-opportunity index: This index is a modification of the Gini coefficient that is used to measure the extent of the inequality between the different type's opportunity sets. Types must be ranked according to the expression $A_j = \mu_j(1 - G_j)$, μ corresponding to the mean outcome and G to the Gini coefficient; the smallest one is the first. Then, the Gini-opportunity index is:$GOp = (1/\mu) \sum_i^k \sum_{i<j} p_i p_j [\mu_j(1 - G_j) - \mu_i(1 - G_i)]$. The result of the computation is the weighted average of the differences between the type's opportunity sets in which the weights are the sample weight of the different types (p_{ij})

[7] They implemented an ordered logit on the five (5) categories of SAH. Three different equations were formulated to measure for direct and indirect effects of circumstances.

Roemer's scheme of IOp[8]. A full information maximum likelihood estimation is applied to the system of equations of the demand for health and for each effort element. The empirical formulation of the model consists of a one-period version of the system of equations in which health outcomes and each of the effort factors depend solely on circumstances and unobserved factors[9].

Following a different methodology, Jusot et al. (2013) analyse the way inequalities of opportunity in health are constructed using multivariate regressions. They analyse the effect of social and family background on the health outcome, controlling for age and sex. To estimate the relative contribution of the aspects that affect inequality, the variance decomposition is computed.[10]

Fajardo (2016) implements a non-linear model for health status with age and gender as controls to obtain direct estimates for IOp. Then, a dissimilarity index[11] is calculated to estimate the probability of reporting a 'good health status', and later decomposed using a Shapley decomposition to obtain the share of each circumstance in the total IOp observed. A Gini-Opportunity index is calculated as an alternative measure.

Carrieri & Jones (2018) apply a decomposition-based approach, first by dividing the population into types, then running separate regressions of health outcomes on effort for each type. To obtain the contribution of circumstances and effort to the level of inequality, they used a regression-based decomposition of the Gini coefficient, and an absolute inequality index using the decomposition of the variance.

Davillas & Jones (2020) does not use the stochastic dominance test to avoid the lack of sufficient observations when building the social types, which is an issue with nonparametric approaches. Therefore, a "mean-based approach" is implemented: it implies the replacement of the outcome of each individual by the mean of the type she belongs to. They use OLS to estimate the contribution of circumstances to health. Individual with same circumstances, have the same predicted health outcome. Then, the mean logarithmic deviation (MLD) is used to measure inequality. The Shapley decomposition is applied to obtain the contribution of each circumstance variable to total IOp. It is worth to mention that this is the methodology that has been used to measure IOp in income and wealth in Mexico as well (Wendelspeiss-Chávez-Juárez, 2014; Vélez-Grajales, et al.,2018; Monroy-Gómez-Franco et al.,2018; Monroy-Gómez-Franco & Corak, 2019).

[8] Rosa Dias, (2010) explains that "the production of health at date t is given by production function, f (Et, Ct, mH), where Et denotes observed effort expended at date t, Ct denotes observed circumstances at date t and mH reflects unobserved factors affecting the production of health. The health stock at any date t+1 is given by the production of health at date t+1 and the depreciated health stock from the previous time period (t), where the depreciation rate (d) is positive and smaller than unity.".

[9] Health is represented by a vector with three components (SAH; long-term illness and disability; mental illness) and effort by a vector composed of three lifestyles (cigarette smoking; weekly consumption of fried food; weekly consumption of alcohol).

[10] The variance decomposition is obtained by the covariance between each source of health and the outcome.

[11] The dissimilarity index is a measure proportional to the absolute distance between the distribution of circumstances among those with high outcomes (i.e., excellent health) and the distribution among those with low outcomes (i.e., poor health).

1.2.3 Existing literature gap

On the one hand, to the best of our knowledge, the theory of EOp has not been applied empirically to measure unfair inequality in *adult health* in Mexico. The studies found in the country aim to measure the extent of IOp in income, wealth, and education, leaving the dimension of health unstudied. This gap represents a major hollow for Mexico for three main reasons. First, evidence has shown that Mexico is one of the most unfairly unequal countries in Latin America and the Caribbean -in terms of income-. These inequalities are persistent over the time, are transferred by generations, and arise from differences in gender, ethnic groups, regions, and others (Wendelspeiss-Chávez-Juárez, 2014; Vélez-Grajales, et al.,2018; Monroy-Gómez-Franco et al., 2018). Second, the prevalence of non-communicable diseases (NCD) is high and has increased in a constant rate over the past decades. The prevalence of these diseases is significantly different for women and men, across the lifespan of people, between regions, and ethnic groups (CONEVAL, 2018). Third, evidence in Mexico show that the socioeconomic background in which individuals are born influences the abilities and opportunities they count with to achieve better outcomes (Campos-Vázquez, 2015).

In other words, these aspects translate into the fact that it is unknow how unfairly unequal is the health status of people in a country characterized by unjust income inequality, with increasing sanitary problems in the population, and with intergenerational transmission of inequalities.

On the other hand, the existing empirical literature that analyzes IOp in adult health mostly focuses on developed countries, with little applications to developing ones (See Appendix 3). Moreover, most of these empirical works analyze subjective measures of health as dependent variables, which might suffer from reporting biases (Jusot et al., 2014). The most recent studies have acknowledged this aspect and measured IOp in objective measures of health (biomarkers) (Carrieri & Jones, 2018; Davillas & Jones, 2020; Carrieri et al., 2020).

Consequently, this Thesis aims to contribute to fill part of that gap. First, this study aims to analyze the extent of IOp in adult health in Mexico. Second, the dependent variables chosen are three objective measures of health (biomarkers) that are used to diagnose and monitor diabetes, and liver and cardiovascular diseases, which represent 41% of total mortality in Mexico (Peters, 1992; Campos-Donato et al. 2019; American Diabetes Association, 2019; INEGI, 2018; CONEVAL, 2018). To the best of our knowledge, this will be the first exercise of IOp using biomarkers in Latin America and the Caribbean. Previous IOp studies have used these measures for the United Kingdom and Indonesia (Jusot et al., 2014; Carrieri & Jones, 2018; Davillas & Jones, 2020; Carrieri et al.,2020). Finally, the data used in the analysis is the main source for monitoring and evaluating health policies in Mexico (INEGI, 2018 & 2019). Therefore, it represents an asset for guiding health policies in the country.

Chapter II

Data and Methodology

This chapter presents the theoretical framework in which this Thesis is based (Section 2.1). Additionally, Section 2.2 explains the political, legal and development frameworks in which the choice of circumstances relies on. Section 2.3 explains the source of the data, and the description and construction of the variables used for the analysis. Then, the method to quantify inequality of opportunity (IOp) in adult health status is explained in Section 2.4, along with the limitations of the approach.

2.1 Conceptual framework

This study is rooted on the ex-ante compensation approach of the Equality of Opportunity theory, building on Roemer's (1998, 2002 & 2016) contributions, and following the literature of IOp in health (Rosa Dias, 2010, Jusot et al., 2013; Carrieri and Jones, 2018; Davillas & Jones, 2020).

To illustrate this framework, let us assume a population composed by individuals represented by i ∈ {1, . . . ,N}, which are characterized by three elements {yi, Ci, Ei }: y denotes a health measure; C is a vector of circumstances and E represents the effort exerted by individuals.

As explained in Section 1.1, circumstances are aspects beyond individual's responsibilities, for which they cannot be held accountable (e.g. ethnicity, sex, age, family's socioeconomic background, the place of birth, and inherited characteristics), and the inequality arising from them is unfair (Roemer, 1998; Ferreira & Peragine, 2015). In contrast, E represents elements under the individual's control and responsibility; therefore, they should be held accountable for and societies should not compensate for them. In the health literature, these aspects are usually related to (un)healthy lifestyles (Rosa-Dias, 2009; Rosa-Dias, 2010; Carrieri & Jones, 2018; Carrieri et al., 2020). However, efforts are influenced by circumstances, but not vice versa. (Roemer, 1998 & 2002; Bourguignon et al., 2007; Checchi & Peragine, 2010; Ferreira & Gignoux, 2011; Carrieri & Jones, 2018; Davillas & Jones 2020).

Then, a health production function is given by:

$$y_i = f(C_i, E(C_i, v_i), u_i) \tag{1}$$

v_i denotes the random variation in effort that is not explained by circumstances, while u_i constitute the random variation in health that is independent to C and E.[12]

Following the ex-ante compensation approach, observing efforts is not a requirement (Ferreira and Gignoux, 2011; Fajardo, 2016; Jusot et al. 2014). However, this approach "does require agreement on some valuation of the opportunity set faced by people" (Ferreira and Gignoux, 2011). Additionally, Roemer (1998) state that the set of circumstances should be decided by societies. Therefore, the choice of variables that reflect conditions beyond people's control is done based on the literature of IOp, and the aspects that the Mexican society has stablish to be unfair. Mexico's national plans, constitution and other official documents were revised for the determination of circumstances, as explained Section 2.2.

After the set of circumstances is defined, one can divide the population into *types* (i. e. group of individuals that share the same circumstances). Let us assume that all circumstances are categorical variables, thus, one can define each *type* by the unique combination of the different circumstances (See Appendix 1 for a detailed explanation). Then, non-parametric approaches can be implemented by applying a measure of inequality to evaluate the extent of IOp *between* the different groups. This approach is very restrictive if one tries to measure the extent of IOp in the population, because the larger the amount of C, the larger the number of subgroups created, leading to insufficient amount of observations to obtain significant results (Roemer, 1998; Ferreira and Gignoux, 2011 & 2014; Davillas & Jones 2020) .

To avoid this issue, the literature proposes measuring the extent of IOp by applying regression-based approaches (Bourguignon, et al., 2007; Ferreira and Gignoux, 2011 & 2014; Paes de Barros et al.,2007; Soloaga et al., 2013; Carrieri and Jones, 2018; Davillas & Jones). Ferreira and Gignoux's (2011) evidenced that the IOp estimates obtained from dividing the population into types and calculating inequality are similar to the ones from the regression-based approach they propose.[13]

Ferreira and Gignoux's (2011) and Davillas & Jones' (2020) methodology is followed: a regression-based approach is implemented, then the MLD to measure the extent of inequality of opportunity in the set of biomarkers chosen. The results of this approach must be interpreted as *lower-bound* estimates of the real IOp since the survey used do not gather data on the full set of circumstances that the literature and the Mexican society have identified as aspects that people should be compensated for.

[12] These two aspects can be thought to represent 'luck' as expressed in Lefranc et. al (2009) and Lefranc & Trannoy (2017).

[13] For a detailed proof, see Ferreira & Gignoux (2011)

2.2 Political, legal and development frameworks of Mexico: defining circumstances.

The choice of circumstance variables is based on the political, legal, and development framework of Mexico, just as recommended by Roemer (1998). Additionally, the choices were also guided by the literature on IOp.

The main framework that guides all policies in the United Mexican States is their Constitution. In this document, the basic principles of the nation are stablished, and the duties and rights of people as well. Therefore, it must be the main framework for the definition of circumstances in this research. First, Article No. 1 prohibits all discrimination that reduces rights and liberties of people due to ethnic background, gender, age, disabilities, social conditions, and others. Second, Article No. 2 states that Federations, States, and municipalities must promote *equality of opportunity* for indigenous people, eliminate any discriminatory practice, and guarantee the access to health services for this population. Article No. 3 stablishes the right of people to receive education, being obligatory for everyone until the level of 'medium superior'. Article No. 4 asserts that women and men are equal, and that every person has the right to protection health (Political Constitution of the United Mexican States, 2020).

Mexico's National Development Plan (PND) 2013-2018 has explicitly stated that inequalities in terms of income, human right violations, discrimination, limited access to health are unacceptable in terms of social justice and represent a challenge to the country. In consequence, both 2007-2012 and 2013-2018 PNDs have included the goals and action plans to achieve 'equality in opportunities'. Special attention is put into reducing inequalities between sexes, throughout the life cycle, indigenous people, and between regions.

Furthermore, the 'Mexican Law of Social Development' 2018 is based on the principle of 'Respect for diversity', which expresses the need to overcome every discrimination practice due to ethnic origin, gender, age, health conditions, religion, etc., and promote an equitable development.

The 'General Law for the Equality between Men and Women' 2018 has the objective to regulate and guarantee equality of opportunities and equal treatment between these two groups, in order to fight against all kinds of discrimination based in the sex of individuals. Furthermore, the Law to Prevent and Eliminate Discrimination has the aim of promoting equality of opportunities and treatment by preventing and eliminating every kind of discrimination. It claims as unacceptable all discrimination due to ethnic or national origin, color of skin, gender, age, and social, economic or health condition, religion, genetic characteristics, and others.

In terms of education, the Law on Education states that everybody must undertake preschool, primary, secondary and middle superior education, and that parents and tutors must send the minors of 18 years to school (General Law on Education, 2019).

As explained before, the Mexican laws, and political and development frameworks pay important attention to the inequalities arising from the differences due to gender, age, ethnic background (putting more emphasis on indigenous people), geographical locations (i.g. regions, urban and rural). Therefore, their inclusion in our analysis is imperative. These variables also coincide with the previous applications of the theory of EOp in health (see Table 1). Additionally, as the level of education until 'medium superior' is mandatory for all inhabitants, it is considered as a circumstance, as in Carrieri & Jones (2018) and Davillas & Jones (2020).

Table 1: Relation between chosen Circumstances, and the empirical literature on inequality of opportunity (IOp) and the Mexican political, social, and legal frameworks

Circumstances	**IOp in health**	**IOp in Mexico**	**Mexican political, social and legal framework**
Sex	Gallardo et. Al (2017), Carrieri & Jones (2018), Davillas & Jones (2020)	Paes de Barro et al. (2009), Wendelspeiss-Chávez-Juárez (2014), Vélez-Grajales, et al (2018), Monroy-Gómez-Franco & Corak, 2019, Soloaga & Wendelspiess (2013), Monroy-Gómez-Franco et al, (2018)	Mexican Constitution, Mexico's National Development Plan (2013-2018), Mexican Law of Social Development (2018)
Age group	Carrieri & Jones (2018), Davillas & Jones (2020)		
Ethnic group or indigenous condition	Jusot et al. (2014), Fajardo (2016), Gallardo, et al., (2017), Davillas & Jones (2020)	Paes de Barro et al. (2009), Wendelspeiss-Chávez-Juárez (2014), Vélez-Grajales, et al (2018), Monroy-Gómez-Franco et al, (2018), Monroy-Gómez-Franco & Corak (2019), Soloaga & Wendelspiess, (2013)	
Educational attaiment	Trannoy et al (2009), Trannoy, et al., (2009), Rosa Dias (2010), Fajardo (2016), Carrieri & Jones (2018), Davillas & Jones (2020)		Mexican Constitution, Law on Education
Region of birth	Fajardo (2016)	Paes de Barro et al.(2009), Monroy-Gómez-Franco & Corak (2019)	Mexican Constitution, Mexico's National Development Plan (2013-2018), Mexican Law of Social Development (2018)
Zone of residence (urban/rural)	Jusot et al. (2014), Fajardo (2016), Gallardo et. Al (2017)	Vélez-Grajales, et al (2018), Monroy-Gómez-Franco et al, (2018), Monroy-Gómez-Franco & Corak, (2019)	Mexican Constitution, Mexico's National Development Plan (2013-2018), Mexican Law of Social Development (2018)
Level of education of parents	Rosa-Dias (2009), Trannoy, et al., (2009), Rosa Dias (2010), Jusot et al. (2014), Fajardo (2016), Gallardo et. Al (2017), Davillas & Jones (2020)	Paes de Barro et al. (2009), Wendelspeiss-Chávez-Juárez (2014), Monroy-Gómez-Franco et al, (2018), Monroy-Gómez-Franco & Corak, (2019), Soloaga & Wendelspiess, (2013)	Law on Education
Indigenous status of mother or father		Vélez-Grajales, et al (2018), Monroy-Gómez-Franco et al, (2018)	Mexican Constitution, Mexico's National Development Plan (2013-2018), Mexican Law of Social Development (2018)
Hereditary health conditions	Rosa Dias (2009), Trannoy et al., (2009), Rosa Dias (2010), Jusot et al. (2014)		Mexican Law of Social Development (2018)

2.3 Data

2.3.1 Description of the survey

The data used corresponds to the last wave of the National Health and Nutrition Survey of Mexico (ENSANUT) 2018. This survey is carried out every six years since 2006 by the National Institute of Public Health (INSP) and the National Institute of Statistics and Geography (INEGI). The ENSANUT 2018 has the aim of gathering information on the conditions of health and nutrition of people who are nationals or foreign residents in Mexico. It includes questions on socioeconomic and demographic aspects of the population, self-reported health status (physical and psychological), healthcare access, risky behaviors, nutrition, etc. This survey is one of the main sources of data to evaluate and monitor individual's health status and healthcare access of the population in Mexico.

The statistical design of the survey is characterized by being probabilistic, stratified, single-stage and by conglomerates, which allows it to have national representativity.[14] The data in the health component can be disaggregated into rural-urban, regional, and the 32 federal entities.

This thesis focuses on the adults of 20 years and older that were selected for the sections on 'blood samples' and for 'anthropometrics and arterial tension'. The demographic and socioeconomic variables were retrieved from the modules Household Characteristics, and the information on their health from the Adult Health module.

It is worth mentioning that the data collection strategy consisted in two visits of two different teams: The Health team and the Team of Specialists. First, the Health team visited the households to apply the questionnaires that only required the reporting of answers of the household members in question.[15] Second, The Team of Specialists (i.e. two people with the profile of nutritionists, and two more with the profile of health or nursing employees) visited the households interviewed by the Health Team and took the anthropometric measures, blood samples and arterial pressure, and applied the questionnaires on nutrition.

The sample size for the health component is 50,654 dwellings. The number of observations for each model varies according to the health outcome studied and the variables chosen i.e. for the blood tests, the sample for adults is 13,490, and for the anthropometrics is 17,474. Table 3 in Section 2.3.2 summarizes the number of observations for each of the models employed, and the total Mexican population represented.

[14] For details in the sampling design of the survey, see Appendix 4.
[15] The questionnaires completed in this visit were: the household characteristics, Adult Health (20 years or older), Teen Health (10 to 19 years old), and Health service users, and others.

2.3.2 Description of variables: Health in adults and Circumstances

2.3.2.1 Dependent variables: health status in adults

To study the IOp in adult health in Mexico, the indicators chosen as dependent variables are the tests (biomarkers) to diagnose the main risk factors and diseases for the deadliest non-communicable chronic diseases in Mexico: cardiovascular diseases (CVD), diabetes mellitus, and liver diseases, which caused 41.0% of the mortality in 2016 (19.9%, 15.4%, and 5.7%, respectively) (CONEVAL, 2018).

First, as hypertension is the main risk factor for CVD, the indicator for this analysis is the systolic blood pressure (SBP). It was chosen because there is strong evidence of its association with the mortality from CVD (Bundy et al., 2017; Weintraub et al., 2015). A person is diagnosed with hypertension if the SBP ≥200mmHg (Campos-Donato et al. 2019).

The criteria to diagnose diabetes is based on the glucose concentrations or its equivalent, the glycated hemoglobin A1C (HbA1C) (American Diabetes Association, 2019). This test represents the average level of glucose in the blood during the last 2 or 3 months, and it is measured as the percentage of glucose in the blood.[16]

The level of albumin in the blood is an indicator of hepatic failure (liver diseases) (Peters, 1996). Albumin is a protein produced by the liver, and it is one of the most important factors in maintaining plasma colloid osmotic pressure[17], in addition to carrying a series of substances such as hormones and vitamins for distribution in the body (INEGI, 2019). When there is a dysfunction in the liver, the production of albumin decreases in proportion to the liver cells that are destroyed[18]; thus, a lower level of albumin in the blood means a worst situation. In average, the concentrations of albumin in the blood is 4.2 g / dL, with a range of 3.5 and 5.0 g / dL (Peters, 1992).

It is important to note the following aspects: first, these biomarkers are objective measures of health, obtained from blood samples, and arterial pressure samples by specialized health workers. This is an advantage because objective measures of health do not have reporting biases, as self-reported and self -assessed measures of health, which may have implications in terms of robustness of the results (Bago d'Uva et al., 2008; Dowd and Zajacova, 2010).

Second, these biomarkers are classified as 'good surrogate measures of mortality risks' by clinical trials and research in health: they "consistently predicts events in the future" (Weintraub et al., 2015)

[16] Usually, diabetes is diagnosed if the HbA1C ≥ 6.5%.

[17] This is the pressure that keeps fluids inside the veins.

[18] Additionally, during acute or chronic malnutrition the level of albumin also decreases

Third, they are continuous variables. Thus, they do not need to be cardinalized to measure inequality, which is usually done in the health inequality literature. The implication of this is that our results should be interpreted as the redistribution of the health burden of preventable conditions (Davillas & Jones 2020).

Fourth, they are ratio scaled: their measurement is unique to its proportional scaling factor, and they have a “true zero value” (Davillas & Jones 2020). The SBP is measured in mmHg, albumin in g / dL, total cholesterol in mg/dL, and HbA1C is a percentage of the concentrations of glucose in the blood (measured in mg/dL).

2.3.2.2 Description of variables: Circumstances

The choice of the circumstances is based in the political, legal and development framework of Mexico, and the empirical literature of IOp in health, as explained in Section 1.2. Here, the construction of the variables is explained. Table 3 presents the main characteristics of the circumstance (independent variables).

Individual’s characteristics

The variable “sex” was obtained from the section “sociodemographic characteristics” of the members of the household interviewed. This variable takes two values: ‘0’ if the member is woman and ‘1’ if the member is man. The words ‘gender’ and ‘sex’, and their categories (men and males, women and females) are used as synonyms to refer to biological condition in which people are born.[19]

The variable “age group” was computed using the age reported by individuals. Five categories were created, corresponding to the following ranges: 20 to 29, 30 to 39, 40 to 49 50 to 59 and above 60.

‘Speak an indigenous language’ was obtained from information on whether the member speaks any indigenous language or dialect. The variable takes values 1 if yes, and 0 if no.

The variable on educational attainment was computed by grouping the last level of education for the people that had only completed until ‘medium superior’, taking as reference the Mexican Law of Education. Four categories were created: ‘No education’, ‘low’, ‘medium’, and ‘upper medium. The regrouping is illustrated in Table 2. Note that for the analysis only individuals with a level of education until medium superior were considered. The ones that had a higher education were not included because higher educational attainment might be considered as effort.

Geographical variables

Place of birth is represented by the region where the person was born. First, the information on the State of birth was retrieved, then the ‘region of birth’ was obtained by grouping the States according to the region

[19] Even though, the author is aware on the conceptual differences between them.

stablished in Conceptual and Sample Design of the survey. The national territory is divided into 4 major regions: North (8 States), Center (12 States), Mexico City (which includes Mexico City and the conurbed municipalities of the State of Mexico), and South (11 states) (INEGI, 2019). The list of federal states and the regions is in Appendix 5. Additionally, the zone of residence represents the current location of the dwelling, whether it is urban or rural.

Familial background

a. *Socioeconomic background*

The information regarding mothers of the individuals was obtained if the person had her mother living on the same household. If that was the case, the respondent had to specify which person from the household was the mother, the education level and the condition of literacy was retrieved for these people.

The level of education of the mother was obtained by grouping the categories of the highest level achieved into 4 categories, taking the Mexican Law of Education as reference: “no education”, “low”, “medium”, “upper medium” and “high”. Table 2 shows which original categories is included in the new ones.

Table 2: Regrouping of categories for educational attainment

New categories	**Original Categories**
No education	No education
Low education	Preschool
	Primary
	Technical or commerce studies with finished primary school
Medium education	Secondary
	Normal basic
	Technical or commerce studies with finished secondary school
Upper medium education	Preparatory or Highschool
	Technical or commerce studies with finished high school
High education	Normal Bachelor’s degree
	Bachelor’s or professional degree
	Master’s degree
	PhD

Hereditary health conditions

The hereditary conditions were obtained by a set of questions included in the module of Adult Health. Respondents had to answer to whether their mother, father or any sibling have or had diabetes or high level of sugar in the blood, hypertension or high arterial tension, high levels of cholesterol or triglycerides. A dummy for each hereditary condition was created: each variable has two categories, 0 if no one has had the condition, 1 if at least one person has presented this condition.

Table 3: Description of the circumstance's variables (independent variables)

Variable	**Type**	**Definition**
Individual's characteristics		
Sex	Dummy	0 = Female 1=Male
Age group	Categorical	1= 20 to 29 years 2= 30 to 39 years 3= 40 to 49 years 4= 50 to 59 years 5= 60 years and higher
Speaks an indigenous language	Dummy	0= No 1= Yes
Educational attainment	Categorical	1= No education 2= Low education 3= Medium education 4= Upper medium education
Geographical variables		
Place of birth		
Region of birth	Categorical	1= North 2= Center 3= Mexico City (metropolitan area) 4= South
Current location		
Zone of residence	Categorical	0= Rural 1= Urban
Familial background		
Socioeconomic background		
Level of education of mother	Categorical	1= No education 2= Low education 3= Medium education 4= Upper medium education 5=High education
Mother speaks an indigenous language	Dummy	0= No 1= Yes
Hereditary health conditions		
Diabetes history	Dummy	0=no family member with diabetes 1= At least a member with diabetes
Hypertension history	Dummy	0=no family member diagnosed with hypertension 1= At least a member with hypertension
High Cholesterol history	Dummy	0=no family member diagnosed with high cholesterol 1= At least a member diagnosed with high cholesterol

2.4 Empirical implementation

Following Ferreira & Gignoux, (2011) and Davillas & Jones (2020), an ex-ante regression-based approach is implemented to measure the contribution of circumstances to our biomarkers. Then the Mean Logarithmic Deviation (MLD) is computed. The MDL belongs to the generalized entropy class inequality measures[20], and it is characterized by being path-independent decomposable and satisfies the Pigou-Dalton transfer axiom[21]. The MLD uses the arithmetic mean as the reference counterfactual. IOp is then obtained by the difference of total inequality and the inequality from a counterfactual distribution that eliminates all the unfair inequality.

In order to illustrate the methodology, let us begin by setting the relation between a health outcome y, circumstances 'C', and efforts 'E'. As explained in Section 2.1, y is a function of C and E, where E depends on C, but not vice versa:

$$y_i = C_i\alpha + E_i\,\beta + u_i \tag{2}$$

$$E_i = HC_i + v_i \tag{3}$$

$\beta$and α are two vectors of coefficients, and H is a matrix of coefficients linking the circumstances variables to the "effort" variables. This matrix allows for the fact that some of the effort variables are clearly affected by circumstances. u_i and $v_i$represent unobserved determinants, sometimes accounted as luck or other random factors (Lefranc, et al., 2009; Lefranc & Trannoy, 2017).

As explained in Ferreira & Gignoux (2011), in order to measure inequality of opportunity one does not need to estimate the previous set of equations, but their reduced form:

$$y_i = C_i(\alpha + \beta H) + v_i\beta + u_i \tag{4}$$

Which can be estimated using OLS as:

$$y_i = C_i\psi + \varepsilon_i \tag{5}$$

Where:

[20] This class of inequality indexes is based on the concept of «entropy». In thermodynamics, entropy is a measure of disorder. When applied to income distributions, entropy (disorder) has the meaning of deviations from perfect equality. In http://www.fao.org/docs/up/easypol/445/theil_index_051en.pdf

[21] The Pigou-Dalton transfer principle state that the inequality measure should change when there is a transfer among individuals from the outcome in question. In the health literature, as applied in Davillas & Jones 2020, this is interpreted as a transfer of the health burden from the unhealthy ones to the healthier. However, the intention of this is not that the worst-off are going to be better by making other people sicker, but to illustrate the extent of inequality in a population, which should be object of policy intervention by reducing the bad health outcomes from worst-off people.

$$\psi = \alpha + \beta H \tag{6}$$

$$\varepsilon_i = v_i \beta + u_i \tag{7}$$

Where:

y_i represents each of the health outcome for each individual

C_i denotes the vector of circumstances variables related to each health outcomes

ψ represents both the direct effect of circumstances in health, and the indirect effect of circumstances through effort.

ε_i denotes the error term.

Table 4 shows the dependent variables (y) and the vector of circumstances (C) related to them.

Then, one can construct the counterfactual distribution which is a 'standardized distribution' given by[22]:

$$\tilde{y}_i = exp\left[\bar{C}_i \hat{\psi} + \hat{\varepsilon}_i\right] \tag{8}$$

$\tilde{y}_i$ denotes the counterfactual advantage level for each individual which is obtained by assigning the average of circumstances across all individuals $\bar{C}_i$. $\hat{\psi}$ denotes the parameter estimate obtained by the OLS regressions from Eq. 7, and $\hat{\varepsilon}_i$ is the variation of unobserved variables.

Then, the inequality measure can be computed. The formula to compute the MLD is:

$$MLD(y) = \frac{1}{N} \sum_{i=1}^{N} log \frac{\mu}{y} \tag{9}$$

$$MLD(\tilde{y}_i) = \frac{1}{N} \sum_{i=1}^{N} log \frac{\mu}{\tilde{y}_i} \tag{10}$$

The absolute level of inequality of opportunity is obtained by the difference of total inequality and the inequality in the counterfactual distribution:

$$\theta_a = MLD(y) - MLD(\tilde{y}) \tag{11}$$

The relative proportion of IOp in our health outcome is given by:

[22] The literature on inequality with this kind of measure sets the requirement of constructing a counterfactual which sometimes relied on 'smooth distributions' -obtained from the distribution of advantages y and a partition of the population into groups, replacing each individual y with the mean of the group she belongs to. However, in order for the 'path independent decomposability axiom' to apply, one must construct a standardized distribution, which assigns the 'grand mean' (the overall mean) to each individual. Doing this, all the between group inequality is eliminated, by rescaling all sub-groups means. Then inequality of opportunity will be the differences between total inequality and the residual (which can be due to effort factors, luck and unobserved circumstances) (Ferreira and Gignoux, 2011)

$$\theta_r = 1 - MLD(\tilde{y})/MLD(y) \qquad (14)$$

It is important to note that the observed C is less than the theoretical vector of the whole set of circumstances that an individual has been endowed with. A "true" value of IOp in health requires that all variables that are not affected by people's choice or control, should be included. However, this is not possible, and is very unlikely to occur, because in practice there are many limitations in terms of data. This is not particular to this methodology or this Thesis. However, the implications for this research is that the value of IOp estimated must be interpreted as a *lower-bound* estimates of inequality due to circumstances in the health indicators chosen (Ferreira & Gignoux, 2011; Wendelspiess & Soloaga, 2014; Davillas & Jones 2020).

Table 4: Independent variables included in the models

	Dependent variables (y)			
Independent variables (C)	**Glycated hemoglobin**	**Albumin in the blood**	**Total cholesterol**	**Systolic Blood Pressure**
Individual's characteristics				
Sex	✓	✓	✓	✓
Age group	✓	✓	✓	✓
Speaks an indigenous language	✓	✓	✓	✓
Educational attainment	✓	✓	✓	✓
Geographical variables				
Place of birth				
Region of birth	✓	✓	✓	✓
Current location				
Zone of residence	✓	✓	✓	✓
Familial background				
Socioeconomic background				
Level of education of mother	✓	✓	✓	✓
Mother speaks an indigenous language	✓	✓	✓	✓
Hereditary health conditions				
Diabetes history	✓			
Hypertension history				✓
High Cholesterol history			✓	✓
Number of observations	1,101	1,127	1,127	1,291
Expanded sample	10,149,915	10,320,437	10,320,437	9,124,738

To measure inequality of opportunity, the STATA module on 'iop' developed by Wendelspiess & Soloaga (2014) was used. Specifically, the 'fg1r' computation implements the regression-based approach and measures inequality using the Mean Logarithmic Deviation as in Ferreira & Gignoux (2011).

2.3.3 Decomposition of inequality of opportunity in health status: Shapley decomposition

The parametric approach implemented provides a point estimate of absolute and relative IOp. However, it is also the aim of this research to calculate the relative importance of each variable into IOp. For this purpose, a Shapley decomposition is applied as in Fajardo (2016) and Davillas & Jones (2020).

The contribution of a circumstance variable is computed by the difference between the total inequality of opportunity and the one that is observed if that circumstance is removed from the set of observed circumstances. The marginal effect of each variable is obtained by estimating the inequality measure for all the permutations of the variables included as circumstances (denoted as k) (Manna & Regoli, 2012; Wendelspiess & Soloaga, 2014). Then, to compute the overall marginal contribution of each circumstance to the overall inequality, one must calculate the average of its marginal effects. Let us denote each circumstance as C_j , j=1, 2,…, k. The contribution of C_j to the explanation of our IOp measure θ is given by:

$$\emptyset\left(C_j,\theta\right) = \frac{1}{k!}\sum_{\pi\in\prod_k}\left[\theta\left(\hat{y}\middle|B(\pi,C_j)\cup\{C_j\}\right) - \theta\left(\hat{y}\middle|B(\pi,C_j)\right)\right] \tag{15}$$

Where $\theta(\hat{y}|X)$ represents the inequality measure, in our case, the MLD; $\prod_k$ denotes the set of all the permutations of the k circumstances; $B(\pi, C_j)$ represents the set of the variables preceding Cj in the ordering π .

The relative share of each circumstance in the total IOp can be written as:

$$s_{j,Shapley} = \frac{\emptyset(C_j,\theta)}{\theta_a} \tag{16}$$

To compute the Shapley decomposition, the STATA module on 'iop' developed by Wendelspiess & Soloaga (2014) was used, with the option 'shapley'.

The advantages of this decomposition consist in that it is path-independent, and the different contributions equal the total value. However, a caveat must me stated: the values obtained from this decomposition approach can only be interpreted as their relative importance, and not as causality (Ferreira & Gignoux, 2014; Wendelspiess & Soloaga, 2014).

2.3.4 Limitations

The main limitations of these methods arise from the availability of the data. First, the survey does not gather information on the full 'true' set of aspects that the Mexican society and the literature on IOp consider to be beyond individual’s responsibilities. Therefore, the results of the parametric approach must be interpreted a *lower-bound* estimates or minimum level of IOp in the health measures included. [23]Thus, the part that is not attributed empirically to IOp, *cannot* be attributed to effort factors or luck.

Second, the measures of adult health status are biomarkers, which indicate the biological state of an individual, used to diagnose a set of chronic diseases in the population. However, they cannot account directly for the multidimensionality of health, as would a multidimensional health index or even a self-assessed health variable (e.g. subjective evaluation of the current health status).

[23] For a detailed proof of this preposition, see Ferreira & Gignoux (2011)

Chapter III

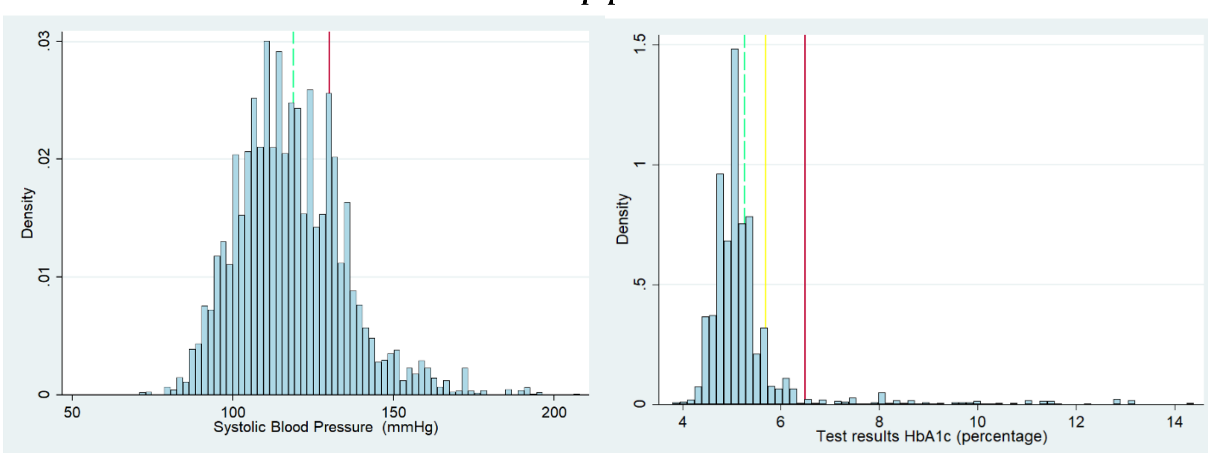
Results

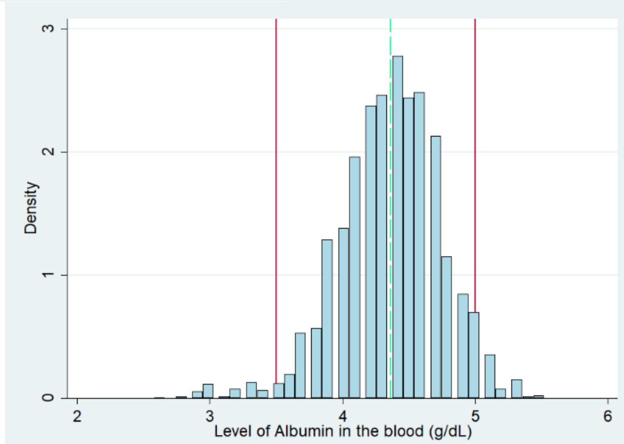
Chapter III presents the descriptive statistics of the population studied (Section 3.1). Then, Section 3.2 presents the point estimates and relative values of Inequality of Opportunity (IOp) for each biomarker for the whole sample, and for the females' and males' samples. Additionally, the Shapley decomposition for each estimation of IOp is presented. Section 3.3 offers a discussion of the results of the empirical exercise, comparing them to previous studies on differences by circumstances on the prevalence of diabetes, hypertension, high cholesterol, and liver diseases in Mexico.

3.1 Descriptive statistics:

Graphs 1, 2, and 3 show the distribution of the biomarkers Systolic Blood Pressure (SBP), albumin in the blood, and glycated hemoglobin (HbA1c). The green lines represent the mean, while the red lines represent the thresholds used to diagnose the chronic diseases they assess. The yellow in Graph 2 represents the limit where a person is diagnosed with prediabetes (American Diabetes Association, 2019).

Graph 1: Distributions of Systolic Blood Pressure (SBP), glycated hemoglobin (HbA1c), albumin in the blood in the studied population





The mean of each biomarkers in the data used for the analysis is within the 'normal' threshold for the diagnose of the disease. [24] However, their distribution is skewed to where it represents a worst health: for the SBP and HbA1c the distributions are skewed to the right. The distribution of albumin in the blood is slightly skewed to the left; lower values of this biomarker is an indicator of hepatic problems in the population (Peters, 1996; INEGI, 2019).

Appendices 6-9 show the descriptive statistics for the biomarkers and circumstances in the studied population. Additionally, the tables contain the estimated means of the biomarker by each circumstance and its categories. On a first glance, one can note the differences in the means between groups, with some common aspects worthy to highlight: older people, women, individuals with lower education levels, and individual's whose mother have a lower education, have means that are closer to the diagnosis of diabetes and hypertension. For albumin levels in the blood, the largest differences are found between sexes, age groups, and level of education of the mother: the groups that have lower means (closer to critical levels) are women, population that was born in Mexico City, middle ages (30 to 49 years), lower education of the mother, and people with low education.

3.2 Inequality of Opportunity in Adult Health Status in Mexico

3.2.1 Absolute and relative estimates for unfair inequality in Adult Health Status

The results presented in this section were obtained as explained in Section 2.3. First, an ex-ante regression-based approach is implemented to measure the contribution of circumstances to our biomarkers. Then, the Mean Logarithmic Deviation (MLD) is used to obtain total inequality and the inequality of a counterfactual distribution. IOp is then obtained by the difference of total inequality and the inequality from the counterfactual. This procedure was implemented for each biomarker in the total samples. Additionally, the models were estimated separately for males and for females to explore how unequal the distribution of biomarkers are within groups. The OLD estimations can be found in the Appendices 10 and 11.

Table 5 shows the *lower-bound* point estimates of absolute and relative IOp in the different biomarkers for the total sample studied, and separate estimations for females and males. As explained in the methodology, these estimates account for the direct contribution of circumstances to inequality, as well as the indirect one, which is due through efforts. It was also stated that the part that is not associated to circumstances, should not be attributed to efforts.

[24] The threshold used in the analysis are based on the following documents: American Diabetes Association, 2019; Campos-Donato et al. 2019; INEGI, 2020; Peters, 1996).

These results show how unequal are the distributions of ‘sugar in the blood’ (glycated hemoglobin - HbA1c), ‘fat in the blood’ (total cholesterol), albumin in the blood, and arterial pressure (systolic blood pressure-SBP) due to observed circumstances. The column (a) contains the point estimates of IOp, and (b) shows the relative contribution of the observed circumstances to total inequality. The bootstrapped standard errors are in parentheses.

Table 5: Absolute and relative inequality of opportunity (IOp) estimates, by health outcome and gender

	Inequality of opportunity (MLD)	
Health measure	**Absolute IOp**	**% of total inequality**
HbA1c Males	0.0032***	**17.20%*** **
	(0.0015)	(0.0015)
HbA1c Females	0.0027***	**17.08%*** **
	(0.0009)	(0.000982)
HbA1c whole sample	0.0028***	**16.57%*** **
	(0.0008)	(0.0008)
Albumin male	0.0004***	**12.02%*** **
	(0.0001)	(0.0001)
Albumin Females	0.0000***	**0.89%*** **
	(0.0001)	(0.0001)
Albumin whole sample	0.0006***	**13.40%*** **
	(0.0001)	(0.0001)
Systolic blood pressure males	0.0004***	**5.31%*** **
	(0.0002)	(0.0002)
Systolic blood pressure females	0.0024***	**23.58%*** **
	(0.0005)	(0.0005)
Systolic blood pressure whole sample	0.0024***	**24.18%*** **
	(0.0004)	(0.0004)

Bootstrap standard errors in parentheses (500 replications)
(Statistical significance levels: + p<0.1 * p<0.05 ** p<0.01 *** p<0.001).

The unfair inequality due to observed circumstances represent significant shares of total inequality for the three health status measures. The *lower-bound* estimates vary across biomarker:13.4% in albumin in the blood, 16.6% for HbA1c, and 24.2% for SBP. These results are in line with the previous application of this for HbA1c and SBP, in Davillas & Jones (2020). There are some differences in terms of the share that they represent, which might be related to the use of different circumstances or differences in inequality levels between the UK and Mexico.

In order to explore how circumstances affect the health status within men and women, separate IOp estimations were computed for each group. The largest difference in terms of the inequality of opportunity that males and females face separately are estimated for albumin in the blood: the minimum unfair inequality in the distribution of this biomarker is 12.0% for males, while for females is 0.9%. The opposite happens for the SBP: the IOp for females is 23.6%, 4.4 times larger than the IOp for males (5.31%), suggesting that the inequality due to differences in circumstances within females are more unequal than within males. The opposite for albumin in the blood, which the unjust inequality is 13.5 times larger for males than for females. The level of IOp for both sexes are almost equal for HbA1c: 17.2% for men, and 17.1% for women.

These results evidence that significant shares of the unequal distributions of ‘sugar in the blood’ (HbA1c), albumin in the blood, and arterial pressure (systolic blood pressure-SBP) are due to unfair aspects that people should not be held accountable for. The separate estimations by gender suggest that unjust inequality affects differently males and females.

The results open the need for a deeper exploration on how each circumstance contribute to the levels of unfair inequality. In the following section, the results from the Shapley decomposition are presented by each biomarker in the total population and for each gender in order to discuss the variations in the contribution of circumstance within females and males.

3.2.2 Decomposition of the inequality of opportunity

Appendices 12, 13 and 14 present the absolute and relative contribution of each circumstance obtained from the Shapley decomposition. Graph 4, 5, and 6 in the main text show the relative contribution of each circumstance to the respective biomarker.

a. *Systolic blood pressure (SBP)*

The IOp from observed circumstances represents almost a quarter of the estimated inequality for SBP. When analyzing the contributions of each circumstance, the importance of sex and age comes to light. The former represents 55.8%, and the latter 27.5%. These results are inverse to the ones obtained by Davillas & Jones (2020), in which the contribution of age was estimated in 61.9%, while for gender 21.9%.

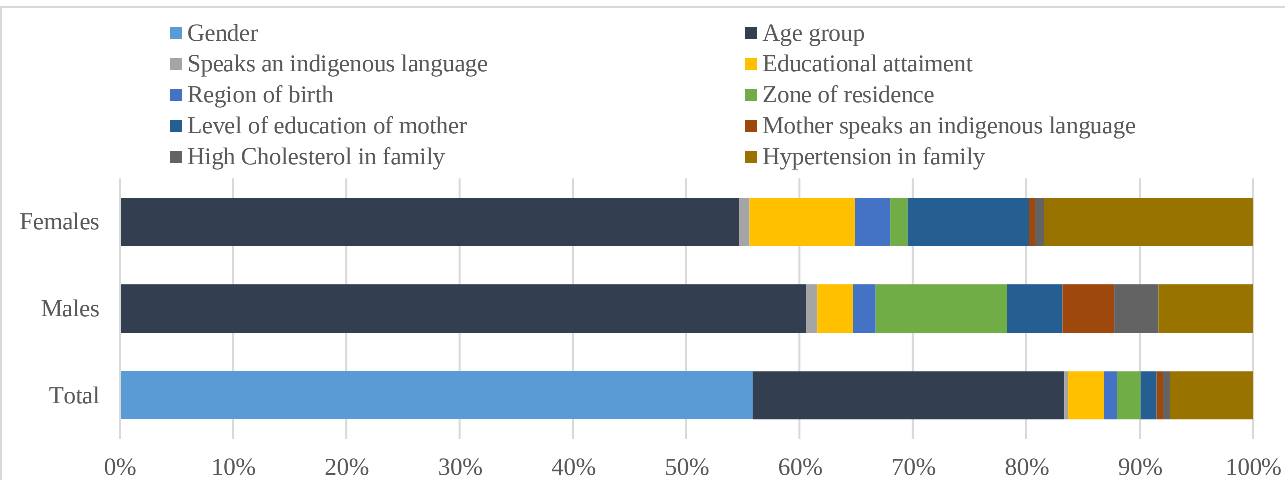
Hypertension history in the family represents 7.3% of IOp, suggesting that variations in genetic factors are associated to differences in the SBP. Furthermore, 3.2% of unjust inequality is associated to the differences in educational attainment.

Even though, for both men and women, IOp is highly associated with age, having or not at least a family member with hypertension is associated with 18.5% of the IOp for females, and for men it represents 8.33%. In contrast, differences in high cholesterol background has a very low association with distribution of SBP in females, but for males is higher 3.9%.

The partial share of the proxy for socioeconomic status in childhood and adolescence (i. e. the level of education of mother), has almost 6 percentage points difference between the shares for females and males. The same difference between these two groups is found for individual's educational attainment (females 9.23% against 3.2% for males).

Current zone of residence has an overall contribution of 2.1%, but for men is up to 11.6%, the second largest share from all circumstances. Lower levels of the observed IOp are associated with the proxies for ethnicity, suggesting that variations between people that speak an indigenous language and those who do not, are not very high. Only for males the share is considerably higher (4.9%).

Graph 2: Decomposition of Inequality of Opportunity (IOp) for Systolic Blood Pressure (SBP)



b. Glycated hemoglobin (HbA1c)

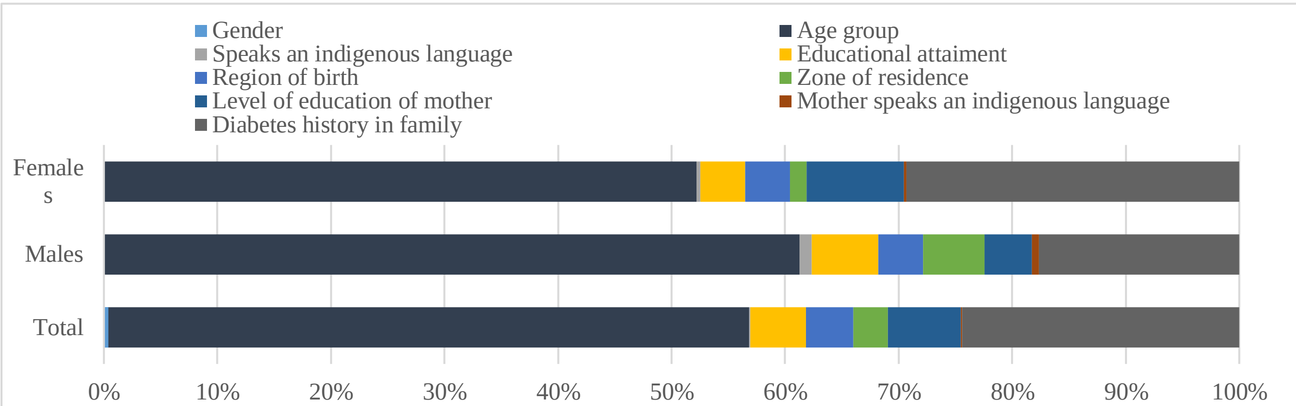
Age is associated with the largest share of all circumstances (56.4%), as in Davillas & Jones (2020), where the largest share of IOp was also associated to age, but to a bigger extent (70.7%).

The variable for diabetes history in family accounts for almost a quarter of IOp for HbA1c in the population. However, one can note that this partial share is larger for females (29.3%), than for males (17.64%). The level of education of the mother is more important than the individual's own education for the overall IOp (6.4%), and for females (8.6%), representing the 3rd largest contributor in both estimations.

Individual's level of education represents almost 5% of the IOp, with similar percentages for the sexes. However, differences arising from speaking or not an indigenous language represent low shares, just reaching 1.0% for males.

Differences in location variables seem to be important for the unfair inequality in HbA1C. For the overall estimation, the region of birth and the current zone of residence together are associated with over 7% of IOp. The partial differences related to these variables are higher for men than for females (9.4% vs. 5.4%).

Graph 3: Decomposition of Inequality of Opportunity (IOp) for glycated hemoglobin (HbA1c)

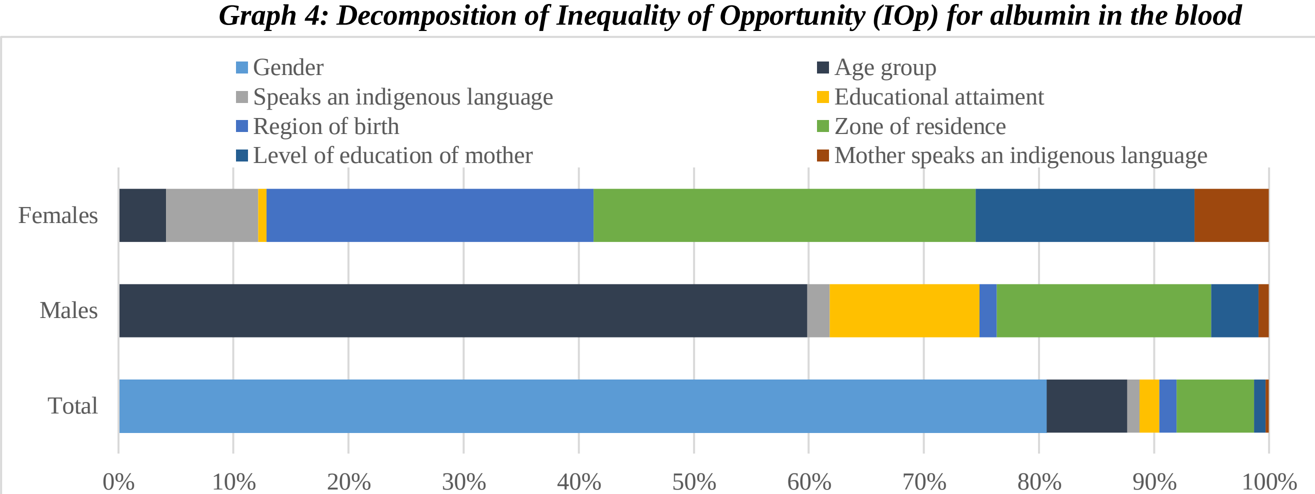


c. Albumin in the blood

An interesting phenomenon happens with this biomarker: gender represents 80.7% of the estimated IOp, which combined with the differences in the separate estimates of IOp for males and females, suggest two main aspects: first, total IOp for albumin in the blood is highly associated to variations between the sexes. Second, when analyzing the estimates separately, one can note that circumstances have a higher share of total inequality for males, implying that differences in these variables are more pronounced within them.

Another differentiating aspect is that the partial proportions associated with each circumstance is very heterogeneous for both sexes. For example, age represents almost 60% of IOp for males, and just 4.2% for females. Individual educational attainment's contribution is 13.0% for males, but 0.7% for females. The educational level of the mother is associated to 19.0% of the unfair inequality in women, while for males is 4.0%.

Graph 4: Decomposition of Inequality of Opportunity (IOp) for albumin in the blood



3.2.3 Discussion

The results presented in the previous sections evidence the existence of Inequality of Opportunity (IOp) in adult health status in Mexico, measured by the levels of sugar in the blood (glycated hemoglobin- HbA1c), albumin in the blood, as well as in the arterial pressure (systolic blood pressure -SBP). Additionally, the extent of the unfair inequality estimates is not trivial, even though they are *lower-bound* estimates. In other words, aspects that individuals are not responsible for affect the distribution of health status associated to the deadliest non-communicable diseases in Mexico. Based on the EOp theory, and the political and social agreements of Mexico, these sources of inequality are unfair, and people should be compensated in order to 'level the playing field'.

In the following paragraphs, the main results of the Shapley decomposition are discussed. Due to lack of studies that analyze IOp in the distribution of these biomarkers, the discussion incorporates researches that study the linkages between the diseases that the biomarkers diagnose (diabetes, liver diseases, hypertension) with the variables included in the analyses. Additionally, this section incorporates studies that investigated how the health and socioeconomic conditions of parents affect the health and other aspects of the lives of individuals.

Age

When analyzing the contribution of the observed circumstances included, the importance of age becomes a highlight for HbA1c, and SBP. In HbA1c age represents the most important circumstance (56.4%), which is in line with the IOp estimations for the UK (Davillas & Jones, 2020). Some studies have evidenced that age is the major risk factor for diabetes (American Diabetes Association, 2019), with a positive direct relation between them (García-García et al., 2002; Cerezo-Correa et al., 2012; WHO, 2014). For hypertension in Mexico, significant differences have been found between the same age groups as in our research, and a positive association between hypertension and age (Campos-Nonato et al., 2019).

However, how much of the differences in age has to do with exclusion and discriminations during the different stages in life? UNDP Mexico (2016) evidenced the persistent levels of inequalities across the lifecycle of people and suggested that these inequalities hinders the opportunities available for individuals. The report states that along the lifespan of people, the exposure to risk factors, and the inequality in earlier stages, accumulate and trigger some diseases later in life, affecting the life expectancy on individuals (UNDP Mexico, 2016). However, the literature revised is not clear on which part of the differences in age is due to the 'normal' aging process, and due to the exposure of risk factors and inequalities during the lifespan. In any case, according to the theory of EOp, inequality arising from age that individual's are not responsible for i.e.

ageing process, and inequalities and exclusions along their lives, should be compensated by public interventions.

Gender

Differences between men and women represent large shares of unfair inequality in Mexico for the distribution of the SBP and albumin in the blood. Studies have evidenced differences in the prevalence and risk exposure of hypertension between men and women: in Colombia, men have 34.0% less the risk to be diagnosed with hypertension than women. (Cerezo-Correa et al., 2012) However, a recent study on vulnerable people in Mexico showed that men have significantly larger prevalence of the disease (Campos-Nonato et al., 2019).

It is worthy of attention the large share of IOp in albumin in the blood associated to differences between the sexes. Evidence shows that the prevalence of liver diseases is higher in males than in females, with a more rapid progression of the illness in male sex. The literature explains that this is related to *social* and *biological* aspects (Roman et al., 2013). On the one hand, males have higher exposure to risk factors that lead them to develop liver diseases, such as contracting viruses (e.g. hepatitis B and C), smoking and drinking, and unhealthy diets. On the other, there is evidence that males have lower production of antioxidants and hormones that help suppress the progression of chronic liver diseases, compared to females. (Roman et al., 2013; Sagnelli, et al., 2017).

The literature exerts that the expusure to risk factors that men have over women is associated to gender roles and stereotypes in Mexico (INMUJERES, 2007 & 2017) . A study on gender stereotypes in young adults in the country evidenced that, first, alcohol and tabaco consumptions are higher for males in more than 7 percentage points. Second, the comsumption was associated to 'machismo' attitudes, i.e. the cultural association to these risk behaviours to men (Chávez-Ayala, et al., 2013).

However, a question arises from these results: which part of the unfair inequality comes from biological factors, and which part arises from the cultural structure of Mexico? Further research in inequality of opportunity in health status should explore aswers to this questions.

Socioeconomic background, geographical variables, and ethnicity

These variables are usually called as 'environmental' variables (Roman et al., 2013; Sagnelli, et al., 2017) or 'structural' variables (Cerezo-Correa et al., 2012; WHO, 2014 & 2018), because they characterize the environment or context in which individuals develop.

In our study, zone of residence, region or birth, and the indigenous condition of the person were found to be associated to important shares of unfair inequality. Another study found significant differences between these groups in the prevalence of hypertension. It explained that such variations are related to differences in lifestyles, eating habits, and the prevalence of other diseases, that are typical to different indigenous groups, and in rural and urban areas (Campos-Nonato et al., 2019). On the other, as stated before, one of the main risk factors to develop liver diseases is alcohol consumption, especially in Mexico where this practice is widespread across the country (Ramos-Lopez et al., 2015). Even though, the literature states that, first, the pattern of alcohol consumption is strongly related to age, and gender, genetic susceptibility in some groups affect drinking patterns across regions, and ethnic groups in Mexico (Roman et al., 2013; Ramos-Lopez et al., 2015).

Moreover, the circumstance of mother's education contributes significantly to shape the opportunities in health status faced by individuals. A study on social mobility in Mexico evidenced that being born into a higher socioeconomic level was associated to better physical development, proxied by the height of people. It was also evidenced that having a higher socioeconomic background was associated to higher weight, which is not a positive outcome, because higher weight is related to health problems, specifically overweight and obesity, which are risk factors for diabetes and CVD (Campos-Vázquez, 2015).

More importantly, that study sheds light on how these circumstances directly affect the outcome of people, and shape the effort levels that people might exert, later impacting the distribution of health status in the population. It was evidenced that the social environment and socioeconomic background in which people were born and grow up, influences the abilities and opportunities they own to achieve better outcomes. This is related to the fact that individuals had differentiated conditions in terms of the school they were sent, lifestyles, family's education and values, levels of stress, eating habits, and others. Additionally, it was evidenced that the aforementioned aspects also impacted the development of differences in cognitive abilities, socioemotional intelligence, and preferences towards risk (Campos-Vázquez, 2015).

Another study evidenced that the socioeconomic conditions in which people are born affect their life expectancy: the worse the conditions, it is expected a lower life expectancy (Moreno-Jaimes, 2017).

A similar phenomenon occurs with the hereditary health conditions. These variables represent more than a quarter of the IOp for glycated hemoglobin, which is the indicator to diagnose diabetes in the population, and more than 10% for systolic blood pressure. This is in line with some evidence of the intergenerational transmission of health conditions in Mexico, proxied by the life expectancy. A study showed that an individual can expect seven additional months of life for every additional year of the life expectance of parents (Moreno-Jaimes, 2017).

Final remarks

In conclusion, the results from this analysis evidence that the observed circumstances affect the unequal distribution of adult health in the systolic blood pressure (SBP), glycated hemoglobin (HbA1c), and albumin in the blood. It was also evidenced that age and gender represent the largest sources of unfair inequality, followed by, in an heterogenous pattern, the individual's and mother's educational level, hereditary health condition, region of birth and zone of residence, and indigenous condition.

After calculating separate regressions for both males and females, the contribution of others circumstance to the inequality within gender varies. This suggests that not only differences between men and women are associated to large shares of unjust inequality in health related to non-communicable diseases, but also the opportunities they face within themselves are unequal.

The results evidences that unequal opportunities in Mexico are present in adult health, not only in income, wealth and education, as it was stated in previous researches. Moreover, these results, related with previous researches, suggest that not only circumstances directly shape the set of opportunities faced by people, but also the levels of efforts they might exert.

Finally, the results have highlighted some aspects. First, the growing sanitary problems in Mexico caused by non-communicable diseases affects differently to people depending on aspects that they are not responsible for. Second, inequalities *between* and *within* men and women, as well as along the lifecycle, are the most pressing ones. Third, this inequality might exacerbate unequal distributions in other dimensions of welfare and be transmitted to future generations. Taking this into account, the need for policies that aim at 'levelling the playing field' in health are imminent in order to give the opportunities to people that they truly deserve.

3.3 Robustness Check:

In order to assess the consistency of the results, two sensitivity analysis were implemented. First, following Davillas & Jones (2020) and Ferreira & Gignoux (2013), a sensitivity analysis was conducted for the inequality measure to alleviate concerns on the robustness of the results. This exercise consisted in applying another measure of inequality, which is the variance share[25], and compare the results with the original ones from the MLD. Appendix 15 shows the estimates for both inequality measures. For SBP and albumin in the blood, the differences are not statistically significant at any level, which means that the IOp estimates are robust. The difference between the two inequality measures for HbA1c is not significant at 1%.

[25] The variance share represents the part of the total variance in the measures of health chosen (biomarkers) that is imputed to circumstances (Davillas & Jones, 2020)

Second, the same model for the HbA1c and SBP were re-run by removing the variables of hereditary health conditions. The estimations of albumin in the blood remained the same because no hereditary health conditions variable was included in the first place. This procedure allowed to evaluate whether the inclusion of more variables increased the level of IOp in each biomarker. Ferreira & Gignoux (2011) explain that as these estimates are lower-bound, the inclusion of more circumstances should increase the point-estimates. Appendix 16 presents the original estimations, and the ones after removing hereditary conditions. The differences between both models for the biomarkers are statistically significant at 1%, meaning that the inclusion of the hereditary conditions add value to our results.

Conclusion and recommendations

Inequality of Opportunity in health arises when the factors that individuals cannot be held responsible for, shape their opportunities to achieve a certain health outcome. These aspects are called 'circumstances' and are associated to gender, age, place of birth, ethnicity, socioeconomic background, hereditary conditions, and others. The inequality arising from circumstances are unfair, and societies should compensate for them under the 'levelling the playing field' objective (Roemer, 1996, 1998, 2002, 2016).

Due to the pressing health challenges that Mexico faces because of non-communicable diseases (NCD), and the differences in their prevalence and mortality between groups, this Thesis aimed at analyzing the level of Inequality of Opportunity in adult health status in Mexico.

Using a regression-based approach and applying the mean logarithmic deviation (MLD) to data of the National Health and Nutrition Survey 2018, the levels of unfair inequality were computed. Then, the Shapley decomposition was implemented to obtain the contribution of each circumstance to the level of IOp. The measures of adult health used in the study are the indicators (biomarkers) that diagnose and monitor the three deadliest NCD's in the country: glycated hemoglobin (HbA1c), albumin in the blood, and systolic blood pressure (SBP). These biomarkers are objective measures of health, and 'good surrogate measures of mortality risks' (consistently predicts health events in the future).

The results provide evidence of unfair inequalities in the levels of sugar in the blood (HbA1c), albumin in the blood, and the SBP. The *lower-bound* estimates of IOp represents 13.4% of total inequality in the distribution of albumin in the blood, 16.6% for HbA1c, and almost a quarter for SBP (24.2%). The results are consistent to the previous application of this methodology in the UK (Davillas & Jones, 2020). Differences in age and gender are associated to the most important shares of IOp in the biomarkers studied. However, the results also evidence the possible presence of intergenerational transmission of health conditions in the population. Additionally, environmental factors (socioeconomic background, zone of residence, region of birth), account for significant shares of unfair inequality. These results suggest that inequalities in the health opportunities faced by people are associated to their socioeconomic background, and the environment they have been exposed to.

After calculating separate regressions for both males and females, the contribution of circumstance to the inequality within gender varies. This suggests that not only differences between men and women are associated to large shares of unjust inequality in health, but also the opportunities they face within themselves are unequal.

Some noteworthy findings from the study include that biological factors from sex, age and hereditary conditions are behind of a large part of the unfair inequality in health status in the country. The 'environment' related to the conditions in which the individuals are born, and the differences in terms of geographical distribution, also associated to important shares of IOp in health status in Mexico. Moreover, these circumstances directly affect the outcome of people, and could also shape the effort levels people exert, later impacting the distribution of health status in the population.

However, there is still the need to identify which part of the unfair inequality related to age comes from the 'natural' ageing process, and from the exposure to exclusions and inequalities throughout the lifespan of people. Both sources of inequality are unfair, but it is necessary to differentiate them in order to design more efficient policies. On a similar note to the analysis by gender, researchers should focus attention on separating the inequality arising from biological and social causes.

From the results of this Study, some recommendations for the Mexican government arise that are worthy to be mentioned. As the Mexican society has established "equality of opportunity" as one of the main goals to achieve, the government should design policies that compensate people for the inequality in health status that arise from aspects that are beyond individual's responsibility. This would require the combination of programs and projects in the different dimensions of welfare, such as education, income and wealth, and health access. The interventions must be designed and implemented considering to the inequalities arising from regions, urban and rural areas, and socioeconomic background of people. Additionally, these policies should be based in a lifecycle approach in order to identify the inequalities that accumulate along the life of individuals. Special attention should be put to the people that are more prone to develop health diseases because of hereditary health conditions. Finally, interventions must be implemented at the national level that aim at eliminating the cultural structure related to gender roles and stereotypes in risky behaviors and unhealthy lifestyles.

All these elements should contribute to 'leveling the playing field' by compensating individuals for the aspects that they should not be held accountable for, and the structural causes that are behind them so everyone would have the same opportunities to play.

References

American Diabetes Association. (2019). Classification and Diagnosis of Diabetes: Standards of MEdical Care in Diabetes. *Diabetes Care*, S13–S28. doi:https://doi.org/10.2337/dc19-S002

Bago d'Uva, T., O'Donnell, O., & Van Doorslaer, E. (2008). Differential health reporting by education level and its impact on the measurement of health inequalities among older Europeans. *Int. J. Epidemiol.*, 1375–1383.

Bourguignon, F., Ferreira , F., & Menéndez, M. (2007). Inequality of Opportunity in Brazil. *Review of Income Wealth, 53*(4), 585–618.

Bundy, J., Li, C., Stuchlik, P., Bu, X., Kelly, T., Mills, K., . . . He, J. (2017). Systolic Blood Pressure Reduction and Risk of Cardiovascular Disease and Mortality: A Systematic Review and Network Meta-analysis. *JAMA Cardiology, 2*(7), 775-781. doi:10.1001/jamacardio.2017.1421

Campos-Nonato, I., Hernández-Barrera, L., Flores-Coria, A., Gómez-Álvarez, E., & Barquera, S. (2019). Prevalence, diagnosis and control of hypertension in Mexican adults with vulnerable condition. Results of the Ensanut 100k. *Salud Publica Mexico, 61*, 888-897.

Campos-Vázquez, R. (2015). *Informe de Movilidad Social 2015: Promoviendo la Movilidad Social en México.* Mexico City: El Colegio de Mexico.

Carrieri, V., & Jones, A. (2018). Inequality of opportunity in Health: a decomposition-based approach. *Health Economics, 27*, 1981-1995.

Carrieri, V., Davillas, A., & Jones, A. (2020). A latent class approach to inequity in health using biomarker data. *Health Economics*, 1-19. doi: https://doi.org/10.1002/hec.4022

Cerezo-Correa, M., Cifuentes-Aguirre, O., Nieto-Murillo, E., & Parra-Sánchez. (2012). Inequalities in morbidity chronic disease based on structural and intermediate determinants. *Revista Gerencia y Políticas de Salud, 11 (23)*, 165-188.

Chávez-Ayala, R., Rivera-Rivera, L., Leyva-López, A., Sánchez-Estrada, M., & Lazcano-Ponce, E. (2013). Orientación al rol de género y uso de tabaco y alcohol en jóvenes de Morelos, México. *Salud pública Méx*, *55*(1).

Checchi, D., & Peragine, V. (2010). Inequality of Opportunity in Italy. *The Journal of Economic Inequality, 8*, pages429–450.

Cohen, G. (2011). *On the Currency of Egalitarian Justice, and Other Essays in Political Philosophy*. (O. Michael, Ed.) Princeton: Princeton University Press.

Consejo Nacional de Evaluación de la Política de Desarrollo Social (CONEVAL). (2018). *Estudio Diagnóstico del Derecho a la Salud 2018* . Mexico City: CONEVAL.

Córdova-Villalobos, J., Barriguete-Meléndez, J., Lara-Esqueda, A., Barquera, S., Rosas-Peralta, M., Hernández-Ávila, M., . . . Aguilar-Salinas, C. (2008). Las enfermedades crónicas no transmisibles en México: sinopsis epidemiológica y prevención integral. *Salud Publica Mex* , 419-427.

Davillas, A., & Jones, A. (2020). Ex ante inequality of opportunity in health, decomposition and distributional analysis of biomarkers. *Journal of Health Economics, 69*.

Dowd, J., & Zajacova, A. (2010). Does self-rates health mean the same thing across socioeconomic groups? Evidence from biomarkers data. *Ann. Epidemiol.*, 743-749.

Fajardo G., J. (2016). Inequality of opportunity in adult health in Colombia. *Journal of Economic Inequality, 14*, 395–416. doi:10.1007/s10888-016-9338-2

Ferreira, F., & Peragine, V. (2015). *Equality of Opportunity: Theory and Evidence.* Bonn, Germany: Institute for the Study of Labor (IZA).

Ferreira, F., & Gignoux, J. (2011). The measurement of inequality of opportunity: theory and an application to Latin America. *Review of Income and Wealth, 57*(4). doi:10.1111/j.1475-4991.2011.00467.x

Fleurbaey, M. (1994). On fair compensation. *Theory and Decision*, *36*, 277-307.

Fleurbaey, M. (1995). Three Solutions for the Compensation Problem. *Journal of Economic Theory, 65*, 505-521.

Fleurbaey, M., & Peragine, V. (2013). Ex Ante Versus Ex Post Equality of Opportunity. *Economica, 80,* 118–130. doi:10.1111/j.1468-0335.2012.00941.x

Fleurbaey, M., & Schokkaert, E. (2008). Unfair inequalities in health and health care. *Journal of Health Economics, 28*, 73–90.

Gallardo, K., Varas, L., & Gallardo, M. (2017). Inequality of opportunity in health: evidence from Chile. *Revista de Saúde Pública*, 51-110.

García-García, E., Aguilar-Salinas, C., Tusié-Lun, T., & Rull-Rodrigo, J. (2002). Early-Onset Type 2 Diabetes in Mexico. *Isr Med Assoc J., 4(6)*, 444-448.

General Law for the Equality between Men and Women. (2018). Diario Oficial de la Federación (DOF 14-06-2018). Retrieved from http://www.diputados.gob.mx/LeyesBiblio/pdf/LGIMH_140618.pdf

General Law on Education. (2019). (2. September 30th, Ed.) Diario Oficial de la Federación.

Government of Mexico. (2007). *Plan Nacional de Desarrollo 2007-2012*. Mexico City. Retrieved from https://www.cenidet.edu.mx/docs/pnd_2007_2012.pdf

Government of Mexico. (2013). *Plan Nacional de Desarrollo 2013-2018*. Mexico City. Retrieved from https://www.snieg.mx/contenidos/espanol/normatividad/MarcoJuridico/PND_2013-2018.pdf

Idler, E., & Benyamini, Y. (n.d.). Self-rated health and mortality: a review of twenty-seven community studies. *Journal of Health and Social Behavior, 38*, 21–37.

INMUJERES. (2017). *Desigualdad en Cifras: Roles y estereotipos de género, una formade discriminación contra las mujeres.* Mexico City: INMUJERES. Retrieved from http://cedoc.inmujeres.gob.mx/documentos_download/BoletinN10_2017.pdf

Instituto Nacional de las Mujeres (INMUJERES). (2007). *El impacto de los estereotipos y los roles de género en México.* Mexico City: INMUJERES. Retrieved from http://cedoc.inmujeres.gob.mx/documentos_download/100893.pdf

Jusot, F., Mage, S., & Menendez, M. (2014). Inequality of Opportunity in Health in Indonesia. *Developpement, Institutions et Mondialisation. Working papers DT/2014–6*.

Jusot, F., Tubeuf, S., & Trannoy, A. (2013). Circunstances and Efforts: How Important is their Correlation for the Measurement of Inequality of Opportunity in Health? *Health Economics, 22*, 1470–1495. doi:10.1002/hec.2896

Kennedy, B., Kawachi, I., Glass, R., & Prothrow-Stith, D. (1998). Income distribution, socioeconomic status, and self rated health in the United States: multilevel analysis. *British Medical Journal*, 917–921.

Law for Social Development. (2018). Mexico City: Diario Oficial de la Federación (DOF 25-06-2018).

Lefranc , A., Pistolesi , N., & Trannoy , A. (2006). Inequality of opportunities vs. inequality of outcomes: Are Western societies all alike? *Society for the Study of Economic Inequality, 54* .

Lefranc, A., & Trannoy, A. (2017). Equality of opportunity, moral hazard and the timing of luck. *Social Schoice and Welfare, 49,* 469–497.

Lefranc, A., Pistolesi, N., & Trannoy, A. (2009). Equality of opportunity and luck: definitions and testable conditions, with an application to income in France. *J. Public Econ., 93,* 1189–1207.

Manna, R., & Regoli, A. (2012). Regression-based approaches for the decomposition of income inequality in Italy, 1998-2008. *Rivista di Statistica Ufficiale*, 5-18.

Monroy-Gómez-Franco , L., Vélez-Grajales , R., & Yalonetzky , G. (2018). *Layers of Inequality: Social Mobility, Inequality of Opportunity and Skin Colour in Mexico.* Mexico City: Centro de Estudios Espinosa Yglesias.

Monroy-Gómez-Franco, & Corak, M. (2019). A Land of Unequal Chances: Social Mobility and Inequality of Opportunity Across Mexican Regions. Retrieved from http://www.ecineq.org/ecineq_paris19/papers_EcineqPSE/paper_446.pdf

Moreno-Jaimes, C. (2017). *Transmisión Intergeneracional de la desigualdad en salud y las políticas públicas en México.* Mexico City: Centro de Estudios Espinosa Yglesias.

National Institute for Statistics and Geography (INEGI). (2018). *Encuesta Nacional de Salud y Nutrición (ENSANUT): Diseño Conceptual.* Aguascalientes: INEGI. Retrieved from https://ensanut.insp.mx/encuestas/ensanut2018/doctos/metodologia/ensanut_2018_diseno_conceptual.pdf

National Institute for Statistics and Geography (INEGI). (2019). *Encuesta Nacional de Salud y Nutrición (ENSANUT): Diseño muestral.* Aguascalientes: INEGI. Retrieved from https://ensanut.insp.mx/encuestas/ensanut2018/doctos/metodologia/ensanut_2018_diseno_muestral.pdf

OECD. (2019). *OECD Economic Surveys: Mexico.* Paris: OECD Publishing. doi:https://doi.org/10.1787/a536d00e-en

Paes de Barros, R., Ferreira, F., Molinas-Vega, J., & Saavedra-Chanduvi , J. (2009). *Measuring Inequality of Opportunities in Latin America and the Caribbean.* Washington DC: World Bank.

Peters, T. (1996). *All About Albumin: Biochemistry, Genetics, and Medical Applications.* San Diego: Academic Press.

Pignataro, G. (2012). Equality of Opportunitu: Policy and Measurement Paradigms. *Journal of Economic Surveys*, *26*(5), 800–834. doi: 10.1111/j.1467-6419.2011.00679.x

Political Constitution of the United Mexican States. (2020). Mexico City: Diario Oficial de la Federación (DOF 06-03-2020). Retrieved from https://www.gob.mx/cms/uploads/attachment/file/547799/1_060320.pdf

Ramos-Lopez, O., Martinez-Lopez, E., Roman, S., Fierro, N., & Panduro, A. (2015). Genetic, metabolic and environmental factors involved in the development of liver cirrhosis in Mexico. *World J Gastroenterol, 21(41)*, 11552-11566.

Roemer, J. (1996). *Egalitarian Perspectives: Essays in Philosophical Economics.* Cambridge: Cambridge University Press. Retrieved from https://books.google.it/books?id=FH-iGlHzr6AC&pg=PP5&source=gbs_selected_pages&cad=3#v=onepage&q&f=false

Roemer, J. (1998). *Equality of Opportunity.* Cambridge, Massachusetts: Harvard University Press.

Roemer, J. (2002). Equality of opportunity: A progress repor. *Social Choice and Welfare*, 455–471. Retrieved from https://www.researchgate.net/publication/24064375_Equality_of_opportunity_A_progress_report

Roemer, J., & Trannoy , A. (2013). Equality of Opportunity. *Cowles Foundation for research in Economics*. Retrieved from https://cowles.yale.edu/sites/default/files/files/pub/d19/d1921.pdf

Rojas-Martínez, R., Basto-Abreu, A., Aguilar-Salinas, C., Zárate-Rojas, E., Villalpando, S., & Barrientos-Gutiérrez, T. (2018). Prevalencia de diabetes por diagnóstico médico previo en México. *Salud Publica Mexico, 60*, 224-232. doi:https://doi.org/10.21149/8566

Roman, S., Zepeda-Carrillo, E., Moreno-Luna, L., & Panduro, A. (2013). Alcoholism and liver disease in Mexico: Genetic and environmental factors. *World J Gastroenterol, 19(44)*, 7972-7982.

Rosa Dias, P. (2009). Inequality of Opportunity in Health: Evidence from a UK Cohort Study. *Health Economics, 18*, 1057–107. doi:10.1002/hec.1535

Rosa Dias, P. (2010). Modelling Opportunity in Health Under Partial Observability of Circumstances. *Health Economics, 19*, 252–264. doi:DOI: 10.1002/hec.1584

Rosa Dias, P., & Jones, A. (2007). Giving Equality Of Opportunity A Fair Innings. *Health Economics, 16*, 109–112. doi: DOI: 10.1002/hec.1207

Sagnelli, E., Stroffolini, T., Sagnelli, C., Pirisi, M., Babudieri, S., Colloredo, G., . . . Almasio, P. (2017). Gender differences in chronic liver diseases in two cohorts of 2001 and 2014 in Italy. *Springer Nature*.

Sen, A. (2002). Why health equity? *Health Economics , 11*, 659–666. Retrieved from https://onlinelibrary.wiley.com/doi/epdf/10.1002/hec.762

Soloaga , I., & Wendelspiess, F. (2010). Desigualdad de oportunidades: aplicaciones al caso de México. In *Movilidad Social en México: Población, desarrollo y crecimiento* (pp. 275-302). Mexico City: Centro de Estudios Espinosa Yglesias.

Trannoy, A., Tubeuf, S., Jusot, F., & Devaux, M. (2009). Inequality of opportunities in health in France a first pass. *Health Economics, 19*, 921-938.

Undén , A.-L., & Elofsson, S. (2006). Do different factors explain self-rated health in men and women? *Gender Medicine, 3*(4), 295-308. doi:10.1016/s1550-8579(06)80218-4

UNDP. (2016). *Informe sobre Desarrollo Humano México.* Mexico City: UNDP Mexico.

United Nations Development Programme (UNDP). (2019). *Human Development Report 2019: Briefing note for Mexico on the 2019 Human Development Report.* New York City: UNDP.

Vélez-Grajales, R., Stabridis-Arana , O., & Minor-Campa, E. (2017). *Still Looking for the Land of Opportunity: The Case of Mexico.* Mexico City: Centro de Estudios Espinosa Yglesias.

Villalpando, S., & Rodrigo, J. (2010). *The status of non-transmissible chronic disease in Mexico based on the National Health and Nutrition Survey 2006.* Mexico City: Salud Pública de México.

Weintraub, W., Luscher, T., & Pocock, S. (2015). The perils of surrogate endpoints. *European Heart Journal*, 2212–2218. doi:doi:10.1093/eurheartj/ehv164

Wendelspiess-Chávez-Juárez , F., & Soloaga , I. (2014). Estimating ex-ante inequality of opportunity. *The Stata Journal, 14*(4), 830–846.

Wendelspiess-Chávez-Juárez, F. (2015). Measuring inequality of opportunity with latent variables. *Journal of Human Development and Capabilities*, 106–121.

WHO. (2013). *Global action plan for the prevention and control of noncommunicable diseases 2013-2020.* Geneva: WHO.

World Health Organisation (WHO). (2014). *Global status report on noncommunicable diseases.* Geneva : WHO.

Appendices

Appendix 1: A sketch model of the indirect approach to equality of opportunity

This section summarizes the 'canonical' model of EOp that Ferreira and Peragine (2015) explained taking the major contributions to the indirect approach in IOp paradigm.

First, let us assume the distribution in the population of an economic good, which can be income, consumption, or an indicator of health outcome. This good is desired by everybody in the population, and there is no limit to that desire. All the determinants of this good, called x from now on, can be identified either as a vector of circumstances (C), which are aspects that individuals are not held accountable for, and effort (E) which are the elements under the responsibility of individuals.

Suppose that circumstances Ω are defined by two variables: gender (male and female), and educational attainment of parents (high school education, college education or higher). The set of combinations of circumstances, or *types*, will be: Ω = ({male, parents with high school education}, {female, parents with high school education}, {male, parents with college education or higher}, {female, parents with college education or higher}. Then, we will have the following equation:

$$x=g(C, E)$$

To simplify, let us assume that the population is characterized by x, C and E, and that each element of the vectors C and e are continuous variables. Then, one can group the population into types and tranches. x_{ij} represent the outcome obtained by C_i and E_j. There are n types, represented by the index i=1,...,n, and m tranches, represented by j=1, ...,m. The population can be organized in a matrix [Xij], in which n rows correspond to types, and m columns, corresponding to tranches.

	e_1	e_2	e_3	…	e_m
C_1	x_{11}	x_{12}	x_{13}	…	x_{1m}
C_2	x_{21}	x_{22}	x_{23}	…	x_{2m}
C_3	x_{31}	x_{32}	x_{33}	…	x_{3m}
…	…	…	…	…	…
C_n	x_{n1}	x_{n2}	x_{n3}	…	x_{nm}

Source: (Ferreira & Peragine, 2015)

To the x_{nm} dimensional matrix [X_{ij}], let there be associated an x_{nm} dimensional matrix [P_{ij}] where each element p_{ij} gives the proportion of total population with circumstances C_i and effort e_j.

Now, let us set the differences the ex-post and ex-ante compensations in the model. On one hand, Ferreira and Peragine 2015, explain that the ex-ante approach aims to assess the set of opportunities which an individual was endowed with, and IOp is eliminated if everyone is endowed with the same set. It specifies that an individual's opportunity set is the series of possible outcomes levels (x_{ij}) that she can obtain given her specific circumstances. In other words, her set of opportunities is composed by all the outcomes of the individuals on her same type. This is represented by the rows of the previous matrix.

On the other, the ex-post approach focuses on assessing the effort exerted by individuals. It investigates the inequalities in groups of individuals that exerted the same degree of effort, or as explained before, tranches (the columns of the previous matrix).

Ferreira and Peragine (2015), argue that measuring IOp using the indirect approach implies a two-step exercise: "first, the actual distribution [X_{ij}] is transformed into a counterfactual distribution [$\tilde{X}ij$] that reflects only and fully the unfair inequality in [X_{ij}], while all the fair inequality is removed. In the second step, a measure of inequality is applied to [$\tilde{X}_{ij}$]" (Ferreira & Peragine, 2015, pp. 18-19). However, the way these two steps are implemented vary significantly depending on the compensation view that researchers take.

In the following paragraphs, the main measurement approaches for both ex-post and ex-ante compensation views are explained.

Ex-ante compensation:

1. The Between-Types Inequality approach: The counterfactual distribution created is obtained by replacing each individual's outcome with the average outcome of the type they belong to. The purpose of this is to eliminate all the inequality within types, as everyone within a type will have the same achievement. Thus, what is evaluated after is the inequality between types.

2. Direct unfairness: The direct unfairness distribution is obtained by replacing each individual's outcome, within a type, with one that would be obtained by a reference effort, given a set of circumstances. The difference compared to the between-types measure is that the direct unfairness takes the achievement of an individual in that type as reference, and then assigns it to everyone in that type.

Ex-post compensation:

1. Within Tranches: This approach replaces each individual's outcome in a tranche with the ratio of her achievement and the average outcome of that tranche. It removes the inequality between tranches and keep same the one within tranches.

2. Unfair inequality: this is also called the Fairness Gap. It replaces each person's income by the ratio of this income and the one that would be obtained by a reference circumstance.

All in all, these measures have the same objective of building a distribution that aims to remove all the "fair" inequality and reflect the unfair inequality in society. After this, an index of inequality can be applied to the resulting distribution.

Appendix 2: Summary of the empirical literature on Inequality of Opportunity (IOp) in Mexico

Authors	Type of survey	Outcome	Circumstance variables	Parametric method	Measure of inequality
Paes de Barro et al. (2009)	cross-sectional	Labor earnings Family income and consumption & PISA results	Gender Ethnicity Birthplace Parents' educational attainment Occupation father	Probit	Mean logarithmic deviation (MLD)
Soloaga & Wendelspiess, 2010	cross-sectional	Level of education and child labor	Education of the household head Access to public services in the household Household's assets ownership Number of siblings	Probit	Dissimilarity Index (DI)
Wendelspeiss-Chávez-Juárez (2014)	cross-sectional	Log income Number of goods at home [0,14] Number of bedrooms per capita Number of goods at home Schooling in years Literacy Current household own a car Quality of the house	Gender Parents' educational attainment Parents owned house SES at age 14 Indigenous condition	OLS	MLD
Vélez-Grajales, et al (2018)	cross-sectional	Income per capita and household assets index	Gender Parents' educational attainment Father's job status Indigenous status Zone of residence (urban/rural) SES status of the household of origin	OLS	MLD
Monroy-Gómez-Franco et al, (2018)	cross-sectional	Household assets index	Parents' educational attainment Father's job status Indigenous status Urban or rural status Household of origin's wealth Current household wealth index Skin tone	OLS	MLD

Authors	Type of survey	Outcome	Circumstance variables	Parametric method	Measure of inequality
Monroy-Gómez-Franco & Corak, 2019	cross-sectional	Household assets index	Sex Skin tone Parents' educational attainment Parents' indigenous condition Current household wealth index Characteristics of the neighborhood of origin Zone of residence (urban/rural)	OLS	MLD
Soloaga & Wendelspiess, 2013	cross-sectional	Level of education	Gender Family log income Father's and mother's years of education Indigenous condition Parent's educational attainment Parent's literacy condition	Probit	DI and Gini Index

Appendix 3: Summary of the empirical literature on Inequality of Opportunity (IOp) in adult health

Authors	Country	Type of survey	Health outcome	Circumstance variables	Effort variables	Method		Measure of inequality
						Non-parametric	Parametric	
Rosa Dias (2009)	UK	Panel Data	Self-Assessed Health (SAH)	Socioeconomic status (SES) background of parents and grandfathers, number of years of schooling of the mother and of the father. Health endowment variables: birthweight, smoking condition of mother during pregnancy, breastfed condition, index of morbidity when child, occurrence of chronic diseases in the parents, diabetes and epilepsy in parents brothers and sisters, obesity in childhood, smoking condition of parents	Cigarette smoking Alcohol consumption Consumption of fried food Educational attainment	FOSD	Ordered probit	Gini-opportunity index Decomposition of the variance
Trannoy A., et al., (2009)	France	Panel Data	SAH	SES background of parents (father's occupation, mother's occupation) Current SES (educational attainment, current or last job) Longevity of parents	Treated effort as a residual influence of circumstances	FOSD	Ordered logit	Gini index
Rosa Dias (2010)	UK	Panel Data	SAH; long-term illness and disability; mental illness)	Parental socioeconomic background, congenital and childhood health, hereditary health conditions, parental smoking conditions, obesity at age 16, cognitive abilities, social development in childhood, educational attainment	Cigarette smoking Alcohol consumption Consumption of fried food Educational attainment	N/a	Ordered probit	
Jusot et al., (2014)	Indonesia	Panel data	A health index (using biomarkers, SRH)	Parent's education, parent's health status, ethnicity, zone of residence, province	Residual of current SES	FOSD	OLS	Variance decomposition

Authors	Country	Type of survey	Health outcome	Circumstance variables	Effort variables	Method		Measure of inequality
						Non-parametric	Parametric	
Fajardo G., (2016)	Colombia	Cross-sections	SAH Health care utilization	Parent's educational attainment, household socio-economic status at age 10, ethnicity, zone of residence, location of birth, region of birth, individual's educational attainment	N/a	FOSD	Logit	Dissimilarity index Gini-opportunity index
Gallardo, et al., (2017)	Chile	Cross-sectional		Mother's educational attainment, family income, zone of residence, sex, region of residence	N/a	FOSD	N/a	
Carrieri and Jones (2018)	UK	Cross-sectional	Cholesterol Glycated hemoglobin Fibrinogen Ill-health index	Cohort of birth, gender, individual's educational attainment until age 18, neighborhood (more vs. less deprived areas)	Saliva cotinine Intensity and frequency of drinking behavior Portions of fruits and vegetables consumed	N/a	N/a	Regression-based decomposition Gini opportunity index Absolute Gini coefficient Decomposition of the variance
Davillas and Jones (2020)	UK	Cross-sectional	Cholesterol Ratio of cholesterol over density lipopı Glycated hemog (HbA1c) C-reactive pı (CRP) Fibrinogen Waist-to-height Systolic blood pre Allostatic load	Gender, age, speaking English at home during childhood, parent's occupation, parent's educational attainment, individual's educational attainment until age 18	N/a	N/a	OLS	Mean logarithmic deviation (MLD). Variance share

Appendix 4: Sampling design of the National Health and Nutrition Survey of Mexico (ENSANUT) 2018

The sample was obtained by, firstly, constructing a stratified set of Primary Sample Units (PSUs) considering the zone of residence, in this case by ‘high urban’, ‘urban complementary’ and ‘rural’(In the political division of the United Mexican States, the localities are divided by their size in high urban (cities with more than 100,000 inhabitants), urban complementary (from 2,500 to 99,999 inhabitants) and rural (with less than 2,500 inhabitants)). Parallelly, the PSUs were grouped into four sociodemographic strata that represent the sociodemographic characteristics of the household members, and the physical characteristics and equipment of the dwellings. [26]Then, the PSUs were selected through a probabilistic sampling, proportional to the size of the dwellings. The informant for the section on household characteristics was the household head or a person of minimum 18 years that knows the information of the household members. The selection of specific informants for the different sections of the survey was through a third stage related to the information on the sociodemographic info on each dwelling (The probability of selecting a person on the k group of interest was obtained by $P^k_{ehij} = P_{ehi} * \frac{1}{Q^k_{ehij}}$; where P_{ehi} is the probability of selection of a dwelling of the i PSU. (INEGI, 2019)

[26] This index was computed using information of the 2010 Dwellings and Population Census, which was the sampling frame for this survey through the Master Sample of the National Housing Framework 2012.

Appendix 5: List of Mexican Regions and Federal States

Region	Federal State	Federal State code
Centro	Aguascalientes	1
	Colima	6
	Guanajuato	11
	Jalisco	14
	México	15
	Michoacán	16
	Morelos	17
	Nayarit	18
	Querétaro	22
	San Luis Potosí	24
	Sinaloa	25
	Zacatecas	32
Mexico City	Ciudad de México	9
Norte	Baja California	2
	Baja California Sur	3
	Coahuila	5
	Chihuahua	8
	Durango	10
	Nuevo León	19
	Sonora	26
	Tamaulipas	28
Sur	Campeche	4
	Chiapas	7
	Guerrero	12
	Hidalgo	13
	Oaxaca	20
	Puebla	21
	Quintana Roo	23
	Tabasco	27
	Tlaxcala	29
	Veracruz	30
	Yucatán	31

Source: National Institute for Statistics and Geography (INEGI) 2019

Appendix 6: Descriptive statistics of the biomarkers and each circumstance

Variable	Obs	Mean	Std. Dev.	Min	Max
Dependent variables (health outcomes)					
HbA1c	15,155,301	5.28	1.09	3.8	14.4
Albumin	15,414,566	4.36	0.38	2.6	5.5
TAS	13,341,626	118.26	16.82	71	208
Independent variables (Circumstances)					
Gender	15,414,566	0.43	0.50	0	1
Zone	15,414,566	0.80	0.40	0	1
Region of birth	15,414,566	2.57	1.08	1	4
Age groups	15,414,566	1.76	1.06	1	5
Educational attainment	10,320,437	3.19	0.80	1	4
Diabetes history	15,414,566	0.39	0.49	0	1
Hypertension history	15,414,566	0.51	0.50	0	1
High Cholesterol history	15,414,566	0.37	0.48	0	1
Mother's indigenous condition	15,414,566	0.06	0.23	0	1
Level of education of mother	15,414,566	2.51	0.99	1	5

Appendix 7: Mean estimations, standard errors, and confidence intervals for systolic blood pressure (SBP), by circumstance

Variable	Mean	Std.Err.	[95%_Conf Interval]	
Individual's characteristics				
Gender			Number of obs= 15,591,553	
Women	129.5322	0.0143503	129.504	129.5603
Man	137.6983	0.0136208	137.6716	137.725
Zone			Number of obs = 15,591,553	
Rural	133.583	0.0115023	133.5605	133.6056
Urban	131.8359	0.0201131	131.7964	131.8753
Region of birth			Number of obs= 15,591,553	
North	138.3618	0.0282527	138.3064	138.4172
Center	136.3549	0.0162856	136.323	136.3868
Mexico City	129.2728	0.0265945	129.2206	129.3249
South	128.1256	0.0158285	128.0946	128.1566
Age groups			Number of obs = 15,591,553	
20 to 29 years	133.5108	0.0138063	133.4838	133.5379
30 to 39 years	126.1939	0.0184524	126.1577	126.23
40 to 49 years	138.3214	0.0278741	138.2668	138.376
50 to 59 years	146.8118	0.0496717	146.7145	146.9092
60 years and higher	138.2053	0.0435708	138.1199	138.2907

Educational attainment			Number of obs = 10,518,518	
No education	138.1746	0.0875559	138.003	138.3462
Low education	139.7453	0.0290397	139.6884	139.8022
Medium education	131.8137	0.018571	131.7773	131.8501
Upper medium education	130.2155	0.0182985	130.1796	130.2514
<u>Familial background</u>				
Mother's indigenous condition			Number of obs = 15,591,553	
No	133.7012	0.0105131	133.6806	133.7218
Yes	125.8724	0.0284466	125.8166	125.9281
Level of education of mother			Number of obs = 15,591,553	
No education	143.4926	0.032211	143.4294	143.5557
Low education	129.7247	0.0138267	129.6976	129.7518
Medium education	133.2426	0.0191343	133.2051	133.2801
Upper medium education	137.5363	0.0376617	137.4625	137.6101
High education	134.1955	0.0409691	134.1152	134.2758
<u>Hereditary health conditions</u>				
Diabetes history			Number of obs = 15,591,553	
No family member with diabetes	134.2493	0.0131838	134.2234	134.2751
At least a member with diabetes	131.5742	0.0153202	131.5442	131.6043
Hypertension history			Number of obs =15,591,553	
No family member diagnosed with hypertension	133.1422	0.0148364	133.1131	133.1713
At least a member with hypertension	133.3116	0.0135854	133.2849	133.3382
High Cholesterol history			Number of obs =15,591,553	
No family member diagnosed with high cholesterol	133.2133	0.0123362	133.1891	133.2375
At least a member diagnosed with high cholesterol	133.2567	0.0171946	133.223	133.2904

Appendix 8: Mean estimations, standard errors, and confidence intervals for albumin in the blood, by circumstance

Variable	Mean	Std.Err.	[95%_Conf Interval]	
Individual's characteristics				
Gender			Number of obs = 15,414,566	
Women	4.245988	0.0001306	4.245732	4.246244
Man	4.516293	0.0001231	4.516052	4.516534
Zone			Number of obs = 15,414,566	
Rural	4.432625	0.0002149	4.432204	4.433046
Urban	4.345747	0.0001087	4.345534	4.34596
Region of birth			Number of obs = 15,414,566	
North	4.362218	0.0002301	4.361767	4.362669
Center	4.397089	0.000159	4.396777	4.3974
Mexico City	4.268951	0.0002595	4.268442	4.269459
South	4.365165	0.0001717	4.364828	4.365501
Age groups			Number of obs = 15,414,566	
20 to 29 years	4.43993	0.0001295	4.439676	4.440184
30 to 39 years	4.257272	0.0002031	4.256874	4.25767
40 to 49 years	4.251279	0.0002318	4.250824	4.251733
50 to 59 years	4.262323	0.0003989	4.261541	4.263105
60 years and higher	4.362993	0.0005037	4.362005	4.36398
Educational attainment			Number of obs = 10,320,437	
No education	4.34914	0.0007896	4.347592	4.350687
Low education	4.269409	0.0002946	4.268831	4.269986
Medium education	4.369866	0.0001897	4.369494	4.370238
Upper medium education	4.369819	0.0002067	4.369414	4.370224
Familial background				
Mother's indigenous condition			Number of obs = 15,414,566	
No	4.361748	0.0001009	4.36155	4.361946
Yes	4.386728	0.0003676	4.386008	4.387449
Level of education of mother			Number of obs = 15,414,566	
No education	4.289067	0.0002805	4.288518	4.289617
Low education	4.356641	0.0001369	4.356372	4.356909
Medium education	4.384327	0.0001984	4.383938	4.384716
Upper medium education	4.426904	0.0003491	4.42622	4.427588
High education	4.371402	0.0003375	4.37074	4.372063
Hereditary health conditions				
Diabetes history			Number of obs = 15,414,566	
No family member with diabetes	4.385831	0.0001235	4.385589	4.386073
At least a member with diabetes	4.327419	0.0001573	4.327111	4.327728
Hypertension history			Number of obs = 15,414,566	

Variable	Mean	Std.Err.	[95%_Conf Interval]	
No family member diagnosed with hypertension	4.378744	0.0001421	4.378465	4.379022
At least a member with hypertension	4.348543	0.0001334	4.348282	4.348804
High Cholesterol history			Number of obs = 15,414,566	
No family member diagnosed with high cholesterol	4.36221	0.0001224	4.36197	4.36245
At least a member diagnosed with high cholesterol	4.364907	0.0001607	4.364592	4.365222

Appendix 9: Mean estimations, standard errors, and confidence intervals for glycated hemoglobin (HbA1c), by circumstance

Variable	Mean	Std.Err.	[95%_Conf Interval]	
Individual's characteristics				
Gender			Number of obs= 15,155,301	
Women	5.320062	0.0003747	5.319328	5.320797
Man	5.221133	0.0004227	5.220305	5.221962
Zone			Number of obs = 15,155,301	
Rural	5.145959	0.0004529	5.145072	5.146847
Urban	5.310272	0.0003317	5.309622	5.310923
Region of birth			Number of obs= 15,155,301	
North	5.135406	0.0004946	5.134437	5.136376
Center	5.17492	0.000383	5.174169	5.175671
Mexico City	5.520481	0.0007547	5.519002	5.52196
South	5.368889	0.0006413	5.367632	5.370146
Age groups			Number of obs = 15,414,566	
20 to 29 years	5.07082	0.0002665	5.070298	5.071342
30 to 39 years	5.249004	0.0004563	5.248109	5.249898
40 to 49 years	5.514263	0.0008676	5.512562	5.515963
50 to 59 years	6.590577	0.0028741	6.584944	6.59621
60 years and higher	6.12566	0.0019967	6.121747	6.129574
Educational attainment			Number of obs = 10,149,915	
No education	5.245004	0.0018491	5.24138	5.248629
Low education	5.792498	0.0013682	5.789816	5.79518
Medium education	5.248657	0.0004875	5.247702	5.249613
Upper medium education	5.185056	0.000499	5.184078	5.186034
Familial background				
Mother's indigenous condition			Number of obs = 15,414,566	
No	5.276908	0.0002881	5.276343	5.277473
Yes	5.281642	0.0012349	5.279222	5.284063

Level of education of mother			Number of obs = 15,155,301	
No education	5.671153	0.0011827	5.668835	5.673471
Low education	5.385914	0.0004652	5.385002	5.386825
Medium education	5.068255	0.0003521	5.067565	5.068945
Upper medium education	5.061709	0.0005232	5.060684	5.062735
High education	5.002953	0.0003837	5.002201	5.003705
Hereditary health conditions				
Diabetes history			Number of obs = 15,155,301	
No family member with diabetes	5.101538	0.000237	5.101074	5.102003
At least a member with diabetes	5.55577	0.0006034	5.554587	5.556952
Hypertension history			Number of obs =15,155,301	
No family member diagnosed with hypertension	5.203886	0.0003827	5.203136	5.204637
At least a member with hypertension	5.346425	0.0004075	5.345626	5.347224
High Cholesterol history			Number of obs =15,155,301	
No family member diagnosed with high cholesterol	5.269057	0.0003547	5.268362	5.269752
At least a member diagnosed with high cholesterol	5.291027	0.0004591	5.290127	5.291927

Appendix 10: Estimated coefficients and standard errors for each model

Circumstances (C)	HbA1c	Albumin	SBP
Gender	0.0140***	0.0621***	0.109***
	(0.0000958)	(0.0000575)	(0.0000831)
Age group	0.0501***	-0.00564***	0.0371***
	(0.0000489)	(0.0000291)	(0.000043)
Speaks an indigenous language	-0.0128***	0.0231***	0.00425***
	(0.000388)	0.000229	0.000302
Educational attainment	-0.00448***	0.00441***	-0.00357***
	(0.0000673)	(0.0000404)	(0.0000588)
Region of birth	0.0135***	-0.00475***	-0.00958***
	(0.0000471)	(0.0000282)	(0.0000409)
Zone of residence	0.0279***	-0.0192***	0.0227***
	(0.000114)	(0.0000682)	(0.0000977)
Level of education of mother	-0.00579***	-9.78e-05**	0.000721***
	(0.0000708)	(0.0000421)	(0.0000626)
Mother speaks an indigenous language	0.00713***	-0.00433***	0.0271***
	(0.000313)	(0.000185)	(0.000245)
Diabetes history in family	0.0673***		
	(0.0000977)		
High Cholesterol in family			-0.0138***
			(0.0000942)
Hypertension in family			0.0365***
			(0.0000908)
Constant	1.499***	1.462***	4.640***
	(0.000316)	(0.000189)	(0.000323)
Observations	10,149,915	10,320,437	9,124,738
R-squared	0.199	0.13	0.244

(Statistical significance levels: + p<0.1 * p<0.05 ** p<0.01 *** p<0.001).

Appendix 11: Estimated coefficients and standard errors for each model

	Health outcomes (y)					
	HbA1c		**Albumin**		**SBP**	
Circumstances (C)	**Females**	**Males**	**Females**	**Males**	**Females**	**Males**
Age group	0.0438*** (0.0000592)	0.0631*** 0.0000845	0.00269*** 0.0000404	-0.0218*** 0.0000391	0.0421*** 0.000055	0.0265*** 0.0000671
Speaks an indigenous language	0.0395*** (0.000518)	-0.0627*** (0.000588)	0.0153*** (0.000344)	0.0270*** (0.000274)	0.0225*** (0.00042)	-0.0254*** (0.00042)
Educational attainment	-6.69E-05 (0.0000858)	-0.00878*** (0.000108)	-0.000327*** (0.0000588)	0.0103*** (0.0000501)	-0.00878*** (0.0000779)	0.00244*** (0.0000867)
Region of birth	0.0113*** (0.0000613)	0.0157*** (0.0000732)	-0.00522*** (0.0000421)	-0.00362*** (0.000034)	-0.0152*** (0.0000559)	-0.00516*** (0.0000581)
Zone of residence	0.0149*** (0.000151)	0.0422*** (0.000173)	-0.0148*** (0.000104)	-0.0250*** (0.0000802)	0.0220*** (0.000134)	0.0272*** (0.000139)
Level of education of mother	-0.0104*** (0.0000958)	0.000859*** (0.000105)	0.00694*** (0.0000654)	-0.0102*** (0.0000487)	-0.0175*** (0.0000877)	0.0141*** (0.0000874)
Mother speaks an indigenous language	-0.0260*** (0.000406)	0.0463*** (0.000489)	0.00891*** (0.00027)	-0.0227*** (0.000228)	0.0184*** (0.000344)	0.0446*** (0.000338)
Diabetes history in family	0.0707*** (0.000127)	0.0633*** (0.000153)				
High Cholesterol in family					-0.0117*** (0.000129)	-0.0146*** (0.000134)
Hypertension in family					0.0494*** (0.000127)	0.0181*** (0.000127)
Constant	1.522*** (0.000404)	1.475*** (0.000493)	1.442*** (0.000277)	1.558*** (0.000227)	4.693*** (0.000432)	4.709*** (0.000477)
Observations	5,725,133	4,424,782	5,837,376	4,483,061	4,905,725	4,219,013
R-squared	0.203	0.208	0.009	0.12	0.239	0.053

Bootstrap standard errors in parentheses (500 replications)
(Statistical significance levels: + p<0.1 * p<0.05 ** p<0.01 *** p<0.001).

Appendix 12: Absolute and percentage of the contribution of each circumstance to inequality (Shapley decomposition) for Systolic Blood Pressure (SBP)

Variable	IOp Systolic blood pressure		IOp SBP Males		IOp SBP Females	
	Absolute	**%**	**Absolute**	**%**	**Absolute**	**%**
Gender	0.1350	55.83%	-	-	-	-
Age group	0.0666	27.53%	0.0321	60.53%	0.1289	54.67%
Speaks an indigenous language	0.0008	0.34%	0.0005	1.00%	0.0022	0.91%
Educational attainment	0.0076	3.15%	0.0017	3.18%	0.0219	9.29%
Region of birth	0.0027	1.13%	0.0010	1.97%	0.0073	3.11%
Zone of residence	0.0050	2.08%	0.0061	11.58%	0.0037	1.57%
Level of education of mother	0.0034	1.41%	0.0026	4.92%	0.0253	10.74%
Mother speaks an indigenous language	0.0014	0.57%	0.0024	4.60%	0.0012	0.49%
High Cholesterol in family	0.0016	0.64%	0.0021	3.88%	0.0018	0.78%
Hypertension in family	0.0177	7.32%	0.0044	8.33%	0.0435	18.45%

Bootstrap standard errors in parentheses (500 replications)
(Statistical significance levels: + p<0.1 * p<0.05 ** p<0.01 *** p<0.001).

Appendix 13: Absolute and percentage of the contribution of each circumstance to inequality (Shapley decomposition) for glycated hemoglobin (HbA1c)

Circumstances	IOp HbA1c		HbA1c Males		HbA1c Females	
	Absolute Iop	**%**	**Absolute Iop**	**%**	**Absolute Iop**	**%**
Gender	0.0007	0.42%	-	-	-	-
Age group	0.0934	56.38%	0.1054	61.28%	0.0891	52.20%
Speaks an indigenous language	0.0002	0.14%	0.0018	1.05%	0.0005	0.31%
Educational attainment	0.0082	4.93%	0.0101	5.88%	0.0068	3.97%
Region of birth	0.0068	4.12%	0.0068	3.98%	0.0067	3.93%
Zone of residence	0.0051	3.08%	0.0093	5.41%	0.0025	1.46%
Level of education of mother	0.0106	6.40%	0.0072	4.16%	0.0147	8.59%
Mother speaks an indigenous language	0.0002	0.11%	0.0011	0.62%	0.0004	0.24%
Diabetes history in family	0.0405	24.43%	0.0303	17.64%	0.0500	29.30%
Total	0.1657	100.00%	0.1720	100.00%	0.1708	100.00%

Bootstrap standard errors in parentheses (500 replications)

(Statistical significance levels: + p<0.1 * p<0.05 ** p<0.01 *** p<0.001).

Appendix 14: Absolute and percentage of the contribution of each circumstance to inequality (Shapley decomposition) for albumin in the blood

Circumstances	IOp Albumin		Albumin Males		Albumin Females	
	Absolute Iop	%	Absolute Iop	%	Absolute Iop	%
Gender	0.1081	80.67%				
Age group	0.0094	7.00%	0.0720	59.88%	0.0004	4.15%
Speaks an indigenous language	0.0015	1.09%	0.0023	1.92%	0.0007	8.01%
Educational attainment	0.0023	1.69%	0.0156	13.01%	0.0001	0.73%
Region of birth	0.0020	1.50%	0.0018	1.51%	0.0025	28.42%
Zone of residence	0.0090	6.74%	0.0224	18.67%	0.0029	33.20%
Level of education of mother	0.0013	1.01%	0.0049	4.09%	0.0017	19.01%
Mother speaks an indigenous language	0.0004	0.30%	0.0011	0.92%	0.0006	6.48%
Total	0.1340	100.00%	0.1202	100.00%	0.0089	100.00%

Bootstrap standard errors in parentheses (500 replications)

(Statistical significance levels: + p<0.1 * p<0.05 ** p<0.01 *** p<0.001).

Appendix 15: Inequality of Opportunity (Iop) measured by the mean logarithmic deviation and the variance

Health measure	Inequality of opportunity (MLD)		Inequality of Opportunity (Variance)
	Absolute Iop	% of Iop	% Iop
HbA1c	0.0028030 (0.0008)	16.57% (0.0008)	19.87% (0.0327)
Albumin	0.0006090 (0.0001)	13.40% (0.0001)	13.00% (0.0334)
Systolic blood pressure	0.002454 (0.0004)	24.18% (0.0004)	24.39% (0.0326)

Bootstrap standard errors in parentheses (500 replications)

Appendix 16: Original estimates of Inequality of Opportunity, and estimates without hereditary conditions

	Original model		**Model 2 (without hereditary conditions)**	
	Inequality of opportunity (MLD)		Inequality of opportunity (MLD)	
Health measure	Absolute Iop	% of total inequality	Absolute Iop	% of total inequality
HbA1c	0.0028030 (0.000808)	16.57% (0.000808)	0.0022800 (0.000726)	13.48% (0.000726)
Systolic blood pressure	0.002454 (0.000415)	24.18% (0.000415)	0.002322 (0.000394)	22.88% (0.000394)

Bootstrap standard errors in parentheses (500 replications)