Essays on the Measurement of Health Inequalities in Mexico

Andrea Salas Ortiz

PhD

University of York

Economics

April, 2022

Abstract

This thesis is about health inequalities in Mexico. The first essay explores the acute obesity crisis in Mexico considering inequality of opportunity. Using John Roemer's framework, *ex-ante* inequality is measured, identified, and characterised in body mass index and waist circumference. Results show that inequalities related to circumstances exist and vary across the whole distribution for both outcomes. Parental health conditions and the geographic region where individuals live are the two main drivers of inequities. The second essay focuses on overcoming the lack of panel data when analysing health inequalities across the lifespan. Using matching and re-weighting methods, a pseudo birth cohort is constructed and the accumulation and intergenerational transmission of *ex-ante* and *ex-post* inequalities in malnutrition are measured. Results indicate that inequities have been persistent across the life course of the birth cohort studied and that lack of opportunities has been transmitted from parents to children. We found consistent evidence pointing out that circumstances are the main driver of inequality in under nutritional outcomes. However, we did not find conclusive evidence that efforts are the main driver of variation in over nutritional outcomes. The third essay investigates the factors behind greater health disparities between indigenous and non-indigenous people in COVID-19 outcomes in Mexico. Using national and administrative public data and making use of a non-linear version of the Oaxaca decomposition method, explained and unexplained differences in hospitalisations and deaths between indigenous and non-indigenous groups are identified and characterised. We find that health disparities are mainly attributable to differences in people's characteristics and that underlying health conditions and household, and municipal socioeconomic characteristics are the main drivers of observable inequalities in hospitalisations and deaths due to COVID-19.

Keywords: Health inequalities; *Ex-ante* and *Ex-post* inequality; Nutritional outcomes; COVID-19; Matching and re-weighting methods; Oaxaca decomposition; Mexico.

JEL codes: D63, I12, I14, I18

To Paulina

List of Contents

List of Figures

List of Tables

Acknowledgements

First and foremost, I would like to thank my supervisors Professor Andrew M. Jones and Professor Nigel Rice. I am grateful for all their support, assistance, input and close supervision throughout my doctoral studies. I am also deeply thankful to Rodrigo Moreno-Serra, my thesis advisory panel member. This work benefited greatly from his thoughtful comments and suggestions.

I dedicate this thesis to Paulina. I cannot thank all the Gods enough for having you in this life. Thank you for being my balance, my supporter, and for keeping my head up and my heart strong. I am as well indebted to life for the parents I have, Elsa and Francisco. I do not think I would be where I am today without all their support. This thesis is also dedicated to my grandfather, Isidro, that passed away two years ago, and always believed in education as the best way to equalise opportunities across generations. I am sure that he and my grandmother, Maria Teresa, would be over the moon to see me, the first woman in the family, to complete a Ph.D. My grandparent's story is just the perfect example of the role of luck, circumstances and efforts in life outcomes. I am also thankful for all the people that have walked with me and made this journey such a wonderful experience. Thanks to my dearest friend, Celia, that introduced me to the best of English right from the beginning. I do not know if I could have made it without scrumptious clotted cream and tangy salt and vinegar crisps! My immense gratitude as well to Georgina Heath and Lenore Klassen, I cannot thank you enough for all the assistance and support throughout these Ph.D. years. Immense gratitude to my closest friends and colleagues, Laia and Chiara. The Ph.D. felt like a much better experience thanks to you. I also feel grateful to my *cuate* Gabriela for being our "bubble" during post-lockdown months. My appreciation also to my pals, Lydia and Alice, for those lockdown nights amid great food and drinks. I am indebted for sharing with me your family recipes for pies, sloe gin and introducing me to Bénédictine. There is a huge *et cetera* of lovely and amazing people that we have met and made York the perfect place for our doctoral studies. We have truly felt the Northern Warmth!

Declaration

I declare that this thesis is a presentation of original work and that I am the sole author of the three chapters. This work has not previously been presented for an award at this or any other University. All sources are acknowledged as References. This work was financially supported by the National Council for Science and Technology in Mexico (CONACYT). I also declare that the funding agencies and data creators have no responsibility for the contents of this thesis. Earlier versions of chapters 1, 2 and 3 are available as working papers (numbers 20/04, 22/07 and 21/20, respectively) in the Health, Econometrics and Data Group (HEDG) Working Paper Series. I have orally presented earlier versions of the three papers at different Seminars at the Department of Economics and Related Studies at the University of York; the 2019 Interdisciplinary Research Network for Economists and Philosophers (IRNEP) conference; the 2019 European Health Economics Association student-supervisor conference; the 2021 International Health Economics Association (iHEA) Congress; the Research Method in Health Economics at National Institute of Public Health in Mexico; the Spanish Economic Association workshop (EVALUES) and the Permanent Seminar on Socioeconomic Inequality at the Colegio de Mexico.

Andrea Salas Ortiz

York, England, April 2022

Introduction

Post-industrial societies defined inequality as a social problem (Bell, 1972). But, in the early Seventies, economists and philosophers questioned whether the concept of inequality inherently implied unfairness. The conception that equality and justice could interchangeably be used as synonyms was firstly challenged by John Rawls. In *Theory of justice*, Rawls introduces the notion of an egalitarian theory of distributive justice. Although Rawls acknowledges that individuals must be responsible for their own life and future, he does not elaborate on the definition of responsibility or to what extent people *can* be responsible (Rawls, 1971). In 1981, Ronald Dworkin introduces a differentiation between legitimate and illegitimate inequality. Dworkin asserts that individuals should not be held responsible for those factors out of their control, but should be accountable for their own choices, decisions, and behaviours. In Dworkin's view, distributive justice must entail a primary and equal allocation of resources. Thus, if all individuals start from an equal point, outcomes should merely be a result of individual's agency (Dworkin, 1981a,b). Nevertheless, Dworkin does not define "individual responsibility" and what it implies in practice. Based on Rawls and Dworkin's work, Roemer (1998) brings a clearer understanding of the term. He agrees about the desirability of an equal playing field for people and asserts that inequality encompasses the differentiation between circumstances (factors beyond people's control) and efforts (acts that reflect individual responsibility) (Roemer, 1998, 2002). Accordingly, Roemer conceptualises equality of opportunity as a situation in which advantages (outcomes) are orthogonal to circumstances. In this regard, equality of opportunity embraces the ethical notion of "responsibility-sensitive egalitarianism".

Motivated by the results shown in *Inequalities in Health: The Black Report* (Townsend P., 1982), health economists have dedicated a great deal of time to explore and propose different methods to measure health disparities. Results systematically indicate that low-income individuals face worse health outcomes. Since the mid of 1980s, greater emphasis has been placed on the measurement of socioeconomic-related inequalities, for example, driven by differences in income, education, household living standards, or a combination of these factors. Rank-dependent indices were the gold standard to measure socioeconomic (SES) health inequalities. Seminal and influential work in this area is the *ECuity project* (1999-2005), which aimed to provide robust evidence about SES inequalities in health and healthcare across European countries. Nevertheless, it became visible that income or the socioeconomic status of a person is partially an illegitimate source of inequality. Inspired by Roemer's *Progress report of equality of opportunity*, the World Health Organisation's 2005 report made explicit that morally objectionable and legitimate inequalities exist. This report set the policy, as well as the research agenda to further explore statistical tools to identify inequality of opportunity (IOp, hereafter).

Conceptually, equality of opportunity (EOp) implies that circumstances at birth should not matter for a person's outcomes later in life but acknowledges the existence of mediating factors that lie on the pathway between circumstances and outcomes. In Roemer's view, the first step in identifying IOp is to divide society into groups, those who share identical circumstances, social *types*, and those who share the same effort, known as *tranches*. Then, two ethical principles emerge: *compensation* and *reward* (Fleurbaey et al., 2013). The former looks for compensation should circumstances play a role in achieving a certain outcome. The latter looks to reward effort exerted by people that experience the same circumstances, but different outcomes. Given these different principles, there have been several approaches to empirically measure IOp. Cecchini et al. (2010) proposed two approaches: the *ex-ante* and *ex-post*. While the *ex-post* approach focuses on outcome differences within individuals with the same level of effort but different circumstances (Ramos et al., 2016), the *ex-ante* approach asserts that equality in opportunities exists if people face the same opportunities before exerting effort or observing any outcome (Roemer et al., 2016). Under the reward principle, there are liberal and utilitarian perspectives. The former claims that if a re-distributive policy exists, this should not be applied to people that share the same set of circumstances but observe different outcomes since these disparities are merely due to differences in efforts. The latter asserts that outcome differences that are explained by differences in efforts should be morally acceptable (Ramos et al., 2016). However, the *ex-ante*/*ex-post* approaches and the liberal/utilitarian reward are potentially in conflict. Fleurbaey et al. (2013) claim that if efforts and circumstances are correlated, the *ex-post* and the *ex-ante* views represent different ways to measure IOp with each representing different ethical principles. If circumstances and efforts are orthogonal and additively separable this incompatibility problem does not exist (Brunori et al., 2022). The empirical relevance of this issue has been further explained by Jusot et al. (2013).

Empirically, many studies have been carried out to measure IOp in health or health care either using parametric or semi-parametric approaches. A comprehensive state-of-the-art overview is presented by Jusot et al. (2019). This text synthesises the literature on the measurement of equality of opportunity in health and healthcare and highlights the approaches followed to measure IOp; the variables that have been used as circumstances and efforts; whether IOp has been measured in the adult or child population, and in which health outcomes IOp has been estimated. In recent years, Jones (2019) set out the theoretical and methodological foundations for a programme of research focused on the measurement of equality of opportunity in health.

Studying inequalities under the IOp framework has the added value of identifying inequities. Inequality and inequity are not synonyms. Equality is about the distribution of a good, such as health, whereas equity is about to what extent there is fairness in the distribution of that good. If there is equity, there are no unfair inequalities. In democratic societies, health inequities are deemed to be socially unacceptable. Thus, the measurement of health inequities in nutrition-related outcomes under the IOp framework permits to disentangles fair and unfair disparities. This is of high relevance for policy implications, especially in Mexico, where obesity and its causes have been framed as a public health problem rooted within the individual sphere and because of people's choices. This has led to public inaction. Hence, this thesis is about the measurement of inequalities in health outcomes in Mexico. Although a particular case, it is suitable to exemplify relevant understudied aspects of the intersection between epidemiological changes with social, economic, and political features that determine health outcomes within a society. Obesity is one clear example. It is probably one of the most challenging global public health problems of our times, and for which no universal solution exists. The Mexican case shows the paradoxical situation in which despite the introduction of several interventions aimed at reducing the rates of obesity (taxes on sugar-sweetened beverages or food labelling), the prevalence of overweight and obesity (OWOB) in the Mexican adult population has continued to rise from 69.7% in 2006 to 75.2% in 2018 (Barquera et al., 2013; National Institute of Public Health, 2018). The narrative around this health problem has been framed, implicitly by some researchers or explicitly by some governments, as *only* a matter of non-constrained individual choices. We argue that this public health problem will not find a solution unless a broader vision that incorporates people's conditions and restrictions on their free will is considered. In this context, Mexico is also an interesting case for studying malnutrition as a continuum that includes both, expressions of overnutrition (overweight and obesity) and instances of undernutrition (stunting, wasting or anaemia). Once more, we argue that it is inaccurate to approach the study of under and overnutrition problems as independent phenomena.

The Mexican case allows us to identify the presence of a double burden of malnutrition and the role that people's circumstances, as well as mediating factors such as people's efforts and choices, have on inequalities that exist, accumulate and transmit across generations. This is an aspect that has been neglected in research settings and the policy-making process alike. Finally, Mexico poses an interesting, yet unfortunate, a case for studying the consequences of the interaction between the COVID-19 pandemic and the malnutrition epidemic, in the context of large socioeconomic disparities and vulnerable indigenous populations.

In the first chapter, *ex-ante* IOp in excess weight proxied by the body mass index (BMI) and excess adiposity measured by the waist circumference (WC), is measured and characterised. The *ex-ante* approach permits us to disentangle observed inequality into a part that can be statistically attributable to exogenous factors such as ethnicity, place of living and parent's health conditions. A lower-bound estimate of IOp is calculated. Given the non-monotonic nature of the outcome, a distributional analysis is also used to look at differences in IOp at the bottom and upper parts of the BMI and WC distributions, where most of the risks to health are found. For example, at the bottom, there is the risk of being malnourished, and at the top percentiles a risk to have moderate and high-risk obesity. To understand which circumstances, matter more to IOp, a decomposition of IOp is undertaken using the Shapley-Shorrocks method.

The second chapter estimates *ex-ante* and *ex-post* levels of IOp in malnutrition outcomes. For this chapter, an age-responsibility threshold approach is considered to differentiate people's circumstances and efforts, such as eating and physical activity patterns and risky health behaviours like smoking or consuming high quantities of alcohol, during childhood, adolescence, and adulthood. Given the lack of panel data that would have allowed us to track individuals over a long period, we rely on two tools used in the evaluation literature: matching and reweighting techniques, to construct a pseudo-birth cohort. We study the accumulation and transmission of IOp in under, overnutrition and an explicit indicator of malnutrition for people born between 1983 and 1988. Since considering *ex-post* IOp, the role of efforts and circumstances is disentangled as well.

The third chapter considers health outcomes resulting from the COVID-19 pandemic. This essay contrasts health disparities between indigenous and non-indigenous people in hospitalisations and deaths in Mexico. The analysis focuses on decomposing ethnic disparities into a part that is attributable to people's observed characteristics and into a part that is not observable or explainable. This is done using an adaptation of the Oaxaca decomposition method that allows for a non-linear relationship between outcomes and covariates.

Results from these essays show that Mexico faces health disparities mainly driven by people's circumstances. Parental and household conditions, as well as the geographic region where individuals live, are factors that boost health disparities. We find that this situation has not changed over 30 years and that inequalities related to circumstances have been transmitted from parents to children and accumulated across time. We do not find conclusive proof that lifestyles or eating patterns play a dominant role in overnutrition outcomes. Finally, we find evidence indicating that COVID-19 has exacerbated pre-existing and longstanding health gaps between indigenous and non-indigenous people.

The measurement of inequalities in health presented in this thesis does not identify causal mechanisms. The analyses are based on normative approaches supported by political philosophy and social choice theory. The main objective is to quantify the extent of social injustice in health. Besides being a valuable approach to differentiate between fair and unfair inequalities, the concept of IOp relies on two ethical principles, equality and freedom and inherently entails a rights-based approach, since some circumstances could explicitly be recognised as protected individual characteristics to any form of discrimination. Equality of opportunity is a condition for a healthier, more productive and fair society (ECLAC, 2017) and represents one of the core elements of liberal democracies that entails the absence of explicit or implicit forms of exclusion to fundamental rights.

Chapter 1

Understanding the role of inequality of opportunity in body mass index and waist circumference among Mexican adults

Abstract. Mexico faces one of the most acute obesity crises worldwide. Despite policy efforts to decrease the prevalence of obesity among adults, an upward trend continues. The aetiology of obesity is complex and defined by multiple causes. While most of the literature has centred on studying behavioural attitudes that contribute to a positive energy unbalance, few studies have explored the role of inequality of opportunity, which focuses on studying the pathways from people's circumstances (factors in which people do not have any control and therefore, cannot be held responsible for) to health outcomes. Using John Roemer's framework, inequality of opportunity is measured, identified, and characterised in body mass index and waist circumference for Mexican adults. Results show that inequalities related to circumstances exist and vary across the whole distribution for both outcomes. Furthermore, parental health conditions and the geographic region where individuals live are the two main drivers of inequality related to circumstances. These findings highlight that to tackle the obesity crisis, a broader approach that accounts for people's circumstances is paramount.

Keywords: Inequality of opportunity in health; distributive justice; inequality related to circumstances; overweight and obesity; Mexico.

1.1 Introduction

Despite the implementation of several health policies and interventions such as the regulation of food and beverage marketing to children, the *National Agreement for a Healthy Nutrition* (Health, 2010), the introduction of new clinical guidelines to diagnose overweight, and the sugar-sweetened beverages (SSBs) tax, Mexico still faces an acute obesity crisis. The prevalence of overweight and obesity (OWOB) in the adult population is the second highest in the world, and has escalated from 69.7% in 2006 to 75.2% in 2018 (Barquera et al., 2013; National Institute of Public Health, 2018). Obesity is a public health problem that represents a health, social and economic burden on a worldwide scale. Obesity decreases the quality and duration of life expectancy (Jarolimova et al., 2013). Furthermore, obesity and its comorbilities jeopardise the sustainability of public health systems and its ability to treat other diseases among the population. The cost of treating obesity amounted to 2.8% of gross domestic product (GDP) of the world in 2014 (Dobbs et al., 2014). For Mexico, it is estimated that 33.2% of the federal public health budget was spent on treating obesity-related comorbilities in 2008. Should the OWOB prevalence continue its rising trend, it is estimated that this cost could increase up to 110%, of the 2008 budget, by 2050 (Rtveladze et al., 2014).

For the sake of simplicity, OWOB is defined as the result of a prolonged positive energy imbalance where energy intake is greater than energy expenditure. However, since obesity is a multiple etiological problem, there are many reasons for this imbalance happening among populations: genes, eating habits, people's living conditions and residency, attitudes and emotions, life habits, income, etc., (Hojjat et al., 2017). The causes of obesity can be classified according to its proximity, as immediate, intermediate, or structural causes that might affect heterogeneously across the life course (Rivera-Dommarco et al., 2018). Immediate causes refer to those factors related to people's lifestyles and behaviours, for instance, high consumption of energy-dense food and/or low physical activity. Intermediate causes are those linked with the production and distribution of food, for example how national food systems evolve. Structural causes are mainly related to the social, economic, and political situation in which people reside.

Mexico offers an interesting case study given its sharp rise in the proportion of its population classified as overweight or obese coupled with its epidemiological and nutritional transition. Many studies have documented the alarming increase in energy intake from SSBs and nonessential high calorific energy-dense food (Barquera et al., 2008; PanAmerican Health Organization, 2015).

Barrientos-Gutierrez et al. (2017) found that the rise in the prevalence of OWOB is due to a greater intake of high-energy food and beverages, as well as changes in lifestyles towards physical inactivity (Barrientos-Gutierrez et al., 2017). Medina et al. (2013) documented an increase in the prevalence of physical inactivity among Mexican adults, while Batis et al. (2018) corroborated that eating patterns of the Mexican population differ substantially from recommendations for healthy living. The interplay between immediate and intermediate causes was studied by Clark et al. (2012) who analysed the effects of the North American Free Trade Agreement (NAFTA) on Mexico's food environment. Their results show that, because of this policy, the Mexican food system has been influenced and modified to the extent that changes in dietary patterns have taken place. Particularly, a higher consumption of soft drinks, snacks, meat, and dairy products among the population (Clark et al., 2012).

Several studies have focused on the link between income and obesity. For example, Pérez Ferrer (2015) found that differences in obesity trends are related to rapid changes in the food environment and cultural institutions, so that people in the lowest deciles of the income distribution have become the most vulnerable to the obesogenic environment (Pérez Ferrer, 2015). Another study found significant associations between socioeconomic indicators such as wealth, education, occupational and marital status, and excess body weight, for both women and men (Quezada et al., 2015). Levasseur (2015) analysed the effect of household socioeconomic status on nutritional outcomes among urban Mexican adults and found strong effects of socioeconomic status on central adiposity, although just for men (Levasseur, 2015). Beltrán-Sánchez et al. (2011) found an association between education and obesity rates. For the case of Mexican men, low education was related to lower obesity, while there was an inverse association for women: more education was related to lower obesity rates (Beltrán-Sánchez et al., 2011). These results indicate that health outcomes and socioeconomic circumstances are related.

The study of socioeconomic inequalities in nutrition-related health outcomes can be undertaken through different approaches. For example, under a pure distributional point of view that focuses on the economic gradient, as in Esposito et al. (2020); Ullmann et al., 2011; Fernald et al. (2007) or using rank dependent indices, as in Clément et al. (2021). However, analysing disparities in nutrition-related outcomes using the IOp framework is useful to unveil prevailing understandings about sources of inequities in nutritional outcomes. This is because the IOp framework explicitly establishes that an individual's health production function involves a combination of circumstances (social determinants of health), efforts (people's choices or free will, decisions) and luck (random shocks). Overall, these have been the conceptual categories used in disciplines like epidemiology or public health when analysing the aetiology of over-nutrition or, overall, nutritional outcomes. Here is an implicit understanding that efforts might potentially be a relevant factor behind nutrition-related outcomes. Indeed, most of the health policies, interventions and programmes that have been implemented in Mexico to tackle the obesity crisis have solely focused on targeting people's behaviours, reflecting the tacit idea that nutrition-related outcomes are mostly related to individual's choices and decisions, although this might not be the case. Indeed, by disentangling the contribution of these health production factors, the IOp framework allows the identification of *inequities* as it disentangles fair and unfair disparities. Equality is about the distribution of a good, such as health, whereas equity is about to what extent there is fairness in the distribution of that good. If there is equity, there are no unfair inequalities. In democratic societies, health inequities are deemed to be socially unacceptable. Thus, the measurement of health inequities in nutrition-related outcomes is of high relevance for policy implications, especially in Mexico, where obesity and its causes have been framed as a public health problem rooted within the individual sphere and as a consequence of people's choices.

The relationship between nutrition-related health outcomes and IOp has already been studied. For example, Nie et al. (2020) focus on the case of China. This study measures IOp in BMI and WC among middle-aged and older Chinese. Results indicate that inequalities are high and range in magnitudes of 65% to 75%, being individual's place of residence its main driver. Research by Bennia et al. (2022) investigates disparities across the BMI distribution in France, the United States of America, and the United Kingdom, stratifying the analysis by sex. They set five different social welfare rankings of BMI categories and use dominance analyses to examine inequalities across the outcome distribution. Their results indicate that BMI is systematically less favourably distributed among women than men. Although the analysis of inequalities is not carried out under the IOp perspective, this study is of special relevance as it considers the ordinal nature of the BMI. Thus, this research contributes to unveiling the existence of unfair health inequalities that are deemed socially unacceptable. Results from this research can be used as robust evidence for policy agenda setting. Furthermore, the identification of the main drivers behind inequities is relevant for targeting purposes. For example, designing interventions that take a family health perspective, rather than individual-focused programmes. Furthermore, it also contributes to the ongoing literature about nutrition-related outcomes and IOp. To the best of our knowledge, it is the first to analyse the case of Mexico.

The application of the IOp framework will allow us to investigate the extent to which opportunities condition people's health outcomes via their choices and behaviours, or the extent to which people had the opportunity to deliberately choose their lifestyles and consumption decisions. Particularly, this study adopts an *ex-ante* approach to identify, measure, and characterise IOp in adults and its role in the OWOB epidemic between 2012 and 2018. The remainder of this paper is as follows: the next section introduces the IOp framework and describes how inequality is measured and decomposed. We measure IOp following a mean-based approach, but given the non-monotonic nature of the outcomes, we use a distributional analysis as well to explore which parts of the distribution circumstances matter the most. The third section describes the data and key variables included in the analysis. Section four presents and interprets the main results, and the final section discusses them within the case of Mexico.

1.2 Definition and measurement of inequality of opportunity

John Roemer (1998) defined two concepts to understand the fairness of (in)equality within a society: circumstances and efforts. Circumstances are situations over which people do not have control and, therefore for which they cannot be held responsible. Sex, ethnicity, parental education, or place of birth are examples of circumstances. Efforts are acts that embrace individual responsibility. For example, life-styles decisions or food consumption choices. IOp is measured following two ethical morals, the compensation, and the reward principles. The former claims that inequalities related to circumstances should be eliminated, and the latter argues to reward efforts among individuals that share the same circumstances. Two approaches to identify inequalities have been suggested: *ex-ante* and *ex-post* (Bruoni, 2016; Ramos et al., 2016). The former is mainly interested in measuring whether, *prior* to exerting any effort or observing any outcome, circumstances are equally distributed. The latter approach looks at what happens *after* efforts and outcomes are observed. The *ex-ante* approach draws on the ethical principle of compensation which claims that if differences in outcomes due to circumstances exist, people should be compensated. From this point of view, equality of opportunity encompasses the ethical position of *responsibility-sensitive egalitarianism* (Roemer, 1998; Roemer et al., 2016) or *levelling the playing field* belief (Jones, 2019). This concept aims to study the pathways from people's circumstances to health outcomes (Jones, 2019) and is applied when concerns about health inequality are tied to questions about access to rights that may guarantee an equal playing field.

This study adopts an *ex-ante* approach to identify, measure, and characterise inequalities related to circumstances and assumes that an equal playing field for people is translated into an equal set of opportunities for everyone, irrespective of whether such opportunities are acted upon. This normative approach allows us to evaluate the extent of individual deviation from a benchmark of *basic* or *minimum* opportunities, to which by law, people are entitled. We make use of parametric models. First, we use a direct and mean-based approach to look at the average level of IOp. Second, since our outcomes are non-monotonic, in the sense that both the bottom and upper parts of the distribution capture illness, we use a distributional approach and measure IOp across the whole distribution. We then capitalise on the linear and additive properties of our model to identify the contribution of each circumstance towards IOp using the Shapley-Shorrocks decomposition approach.

1.2.1 Measuring *ex-ante* inequality of opportunity

We use a parametric and direct approach to measure absolute and relative IOp. *Ex-ante* IOp relies on the idea that if all individuals have the same set of opportunities, circumstances should not be related to outcomes. The direct approach consists of measuring inequalities by relying on a counterfactual of the outcome distribution (Jusot et al., 2019). Thus, we evaluate the extent to which each individual deviates from the *social opportunity set*, which is defined as the average level of advantages across the population (Ferreira et al., 2011). This approach relies on a health production function in which the health outcome, y, of individual i is a function of demographics (sex and age), X, their circumstances, C, their efforts, E, and other random factors, u, such as luck[1], or other situations that the individual cannot avoid. It is relevant to add the three normative assumptions embedded in the IOp framework (Jones, 2019):

- *Responsibility cut*: a partitioning of circumstances and effort is possible and desirable;
- *Linearity*: conditional on circumstances, there is a linear relationship between effort and outcome, and
- *Control*: efforts are a function of an individual's circumstances, but it cannot be assumed that circumstances are a function of people's efforts, since, by definition, circumstances are uncontrollable factors by the individuals.

The model assumes that efforts are in turn determined also by demographics factors, X, circumstances, C, and other random factors that are not people's circumstances, but affect the level of effort exerted, v. The correlation between effort and circumstances can be exemplified by the

[1] Lefranc et al. (2009) define different types of "luck", for example, genetic luck, brute luck, initial and later brute luck.

fact that individuals could claim that they do not exercise in an open-air space because there is an excess of pollution or due to security concerns. Thus:

$$y_i = h(C_i, E_i, X_i, u_i) \tag{1.1}$$

$$E_i = g(C_i, X_i, v_i) \tag{1.2}$$

Assuming additive separability and linearity in $h(.)$ and $g(.)$ a system of equations can be written as:

$$y_i = \alpha_0 + \alpha_1 C_i + \alpha_2 E_i + \alpha_3 X_i + u_i \tag{1.3}$$

$$E_i = \delta_0 + \delta_1 C_i + \delta_2 X_i + v_i \tag{1.4}$$

where α_1 and α_2 are parameters that reflect the direct effect of circumstances and efforts on the outcome, respectively. δ_1 is a vector of coefficients that captures the indirect effect of circumstances on efforts. Then, by inserting Equation (1.4) into (1.3), we have:

$$y_i = (\alpha_0 + \alpha_2\delta_0) + (\alpha_1 + \alpha_2\delta_1)C_i + (\alpha_2\delta_2 + \alpha_3)X_i + (\alpha_2 v_i + u_i) \tag{1.5}$$

and arranging the terms, we get the following linear reduced form:

$$y_i = \beta_0 + \beta_1 C_i + \beta_2 X_i + \epsilon_i \tag{1.6}$$

Here, y_i is the health outcome (body mass index (BMI) or waist circumference (WC)) for individual $i, ..., N$; $\beta_0 = (\alpha_0 + \alpha_2\delta_0)$ is the intercept; $\beta_1 = (\alpha_1 + \alpha_2\delta_1)$ captures the total contribution of circumstances, reflecting the direct effects of circumstances on the outcomes and the indirect effect of circumstances through efforts and, $\epsilon_i = (\alpha_2 v_i + u_i)$ is the error term that captures random variation in outcomes. We estimate Equation (1.6) using Ordinary Least Squares (OLS) accounting for age and sex fixed effects[2]. It is worth noting that, in practice, efforts are not observed or needed in the estimation of Equation (1.6). According to the health production function embedded in the IOp in the health model, it is not possible that circumstances can be determined by people's efforts. For example, people's ethnicity cannot be determined by efforts or parents' formal education cannot de-

[2]How to treat or define age and sex is a matter of debate from a philosophical point of view. Some analyses have included them as circumstances (Davillas et al., 2020a; Ding et al., 2021), others have defined them simply as demographic factors, but not included them in the IOp measurement (Jusot et al., 2013; Nie et al., 2020). In cross-sectional analysis, the inclusion of age as a circumstance reflects the birth cohort each person belongs to. Health inequalities between two individuals of different ages say 31 and 56 years old, reflect technological developments. Despite the urge to compensate the 56-year-old individual for not being exposed to the same technological advances as the 31-year-old, the counterargument is that these situations are unavoidable and something that everyone faces across the life cycle (Jusot et al., 2013). Hence, we took sex and age as fixed control variables.

pend on individuals' choices. In this context, circumstances are exogenous factors to the individual.

This direct and parametric *ex-ante* approach assumes that all individuals have the same opportunities if, in expectation, no differences in outcomes arise because of having different circumstances (Roemer et al., 2016). This expectation over outcomes within individuals who share the same circumstances[3] can be approximated using the mean. This implies the ethical assumption of inequality neutrality, which is in line with the "utilitarian reward" principle (Jones et al., 2014). Thus, the presence of IOp is assessed by comparing the deviation of the actual outcome with the predicted distribution of outcomes, $\mathbb{E}(y_i \mid C_i)$, also known as the *smoothed* distribution (Ferreira et al., 2011)). Thus, with age and sex fixed effects being absorbed:

$$\widehat{y_i} = \widehat{\beta_1} C_i \tag{1.7}$$

In this mean-based approach, the counterfactual outcome is inserted into an inequality measure. The choice of the inequality measure is mainly based on some desirable properties we expect the measure to have. First, it is desirable that the measure meets the basic axioms of normalisation; scale invariance and population replication, but the measure should also satisfy the within-type transfer insensitivity and between-type transfer sensitivity principles. Moreover, since we apply a Shapley-Shorrocks decomposition, it is also necessary for the measure to have the property of additive decomposability. It is well known that only inequality measures that belong to the generalised entropy family class meet these requirements. We use the mean logarithmic deviation (MLD), proposed by Cecchini et al. (2010), as our preferred measure given its path-independent decomposability property. It is worth noting that the within-type transfer insensitivity implies that inequalities among people that share the same set of circumstances do not matter (this is equivalent to assuming inequality neutrality), the only inequality that is relevant under this approach is that associated with opportunities, that is, the inequalities that occur across people with different opportunity sets (Ferreira et al., 2011). The MLD is the generalised entropy index $GE(\alpha)$ when $\alpha = 0$ and is calculated as follows:

$$MLD(\hat{y}) = \frac{1}{N}\sum_{i=1}^{N} ln\frac{\bar{\hat{y}}}{\hat{y}_i} \tag{1.8}$$

Thus, absolute inequality is defined as the MLD of the counterfactual distribution of health out-

[3]It is worth noting that the predicted outcomes will be the same for all individuals who have identical circumstances

comes conditioned on circumstances, such that:

$$\theta_a = I_0(\hat{y_i}) \tag{1.9}$$

where I_0 denotes the MLD, and $\hat{y}$ depicts the counterfactual outcome $\mathbb{E}(y_i|Ci)$. The absolute inequality measures the deviation of the expected level of health outcome from the group's expected average, so if θ_a is zero there is no inequality, and larger values reflect higher levels of inequality.

A relative measure of IOp is estimated by obtaining the ratio of the absolute level of inequality with respect to the overall inequality, as:

$$\theta_r = \frac{I_0(\hat{y_i})}{I_0(y_i)} \tag{1.10}$$

In Equation (1.10), θ_r is defined as the MLD of the counterfactual distribution of outcomes divided by the MLD of the actual distribution of outcomes, the latter defined as overall inequality. Relative inequality is zero when equality is observed, and positive values depict an unequal distribution of the outcomes.

1.2.2 Drivers of *ex-ante* inequality of opportunity

To design policies that address inequalities, the identification of the main drivers behind IOp is paramount. For this, estimates of IOp in BMI and WC are decomposed into their sources and the relative importance of each circumstance to the overall IOp is quantified. This is done using the Shapley-Shorrocks decomposition method. As stated earlier, Equation (1.6) depicts a linear and decomposable model. Thus, we calculate the marginal contribution of each circumstance included in C to the variance in our outcomes. This is calculated as the difference in the variance explained when the c^{th} circumstance is included and the variance when that circumstance is excluded. Differences are calculated using all possible permutations of circumstances. Then, the sum of the differences is averaged across the number of all possible permutations. It is worth noting that the contribution of each circumstance is not equivalent to the casual effect of each circumstance on IOp since unobservable determinants of nutrition-related outcomes are likely to be correlated with the observable circumstances as discussed by Ferreira et al. (2011). Thus, this decomposition only indicates the importance of each circumstance (Chávez -Juárez et al., 2014).

1.2.3 Going beyond the mean: IOP across outcomes distribution

In Equation (1.6), inequality is calculated using the expected conditional mean of the outcome, given the whole distribution of the set of circumstances. This implies assigning homogeneous weights to the contribution of circumstances within those individuals that share the same circumstances, which translates to assuming *inequality neutrality*. Nevertheless, this might not be the case and could be too restrictive for the purpose of uncovering inequalities in OWOB. It is restrictive because by using the mean-based approach, it is not possible to observe the contribution of circumstances in the upper tails of the BMI and WC distribution, where excess weight and adiposity are observed. BMI and WC outcomes are nonlinear health outcomes, lower and higher values of these measures reflect illness in the form of undernourishment or overnutrition. To address this issue, we now assume *inequality aversion*, allowing an allocation of different weights to the contribution of circumstances across the outcome distribution for those individuals that share the same circumstances. We follow the approach used by Davillas et al. (2020a), based on unconditional quantile regressions (UQR). This approach is based on the idea that the outcome in Equation (1.6) can be replaced by a RIF, defined as (Borgen, 2016; Firpo et al., 2009):

$$RIF(y; q_\tau, F_Y) = q_\tau + \frac{\tau - 1\{Y \leq q_\tau\}}{f_y(q_\tau)} \tag{1.11}$$

where q_τ is the value of y at the τ percentile. y in our case is BMI or WC. F_y is the cumulative distribution function of y, and $f_y(q_\tau)$ is the density of y at the q_τ. $1\{y \leq q_\tau\}$ is the indicator function and identifies whether y is below or above the observe quantile q_τ. Using this method, the RIF is the new outcome variable and it can be estimated as:

$$RIF(y; q_\tau) = C_i\beta^\tau + \epsilon_i^\tau \tag{1.12}$$

Where τ=25^{th}, 50^{th}, 75^{th} and 95^{th} percentiles. The two final percentiles capture the top end of the BMI and WC distributions, where excess weight and adiposity are observed. ϵ_i^τ represents the model error. The counterfactuals to be used are proxied by:

$$\widehat{y_i}^\tau = \widehat{\beta^\tau} \tag{1.13}$$

We used the MLD again to measure absolute and relative IOp and likewise exploit the linear and additive properties of the RIF equations to identify the main sources of inequalities using the Shapley decomposition method. Throughout the analysis ENSANUT sample weights are used, which make the results representative of the Mexican population for the two years studied.

1.3 Data

Data from the cross-sectional National Surveys of Health and Nutrition (ENSANUT, using its Spanish acronym) for 2012 and 2018 are analysed. The datasets are nationally representative surveys whose target population are the inhabitants of private households in Mexico. These national cross-sections are multi-stage and stratified surveys (by urbanity and geographical areas). The sample design of the waves permits inferences about the health of the Mexican population. A full and detailed description of the sampling methodology is found elsewhere (Romero-Martínez et al., 2013, 2019).

The datasets consist of a collection of demographics, social and economic conditions, as well as nutrition-related health outcomes of the population, via anthropometric measurements such as weight, waist circumference and height. Even though ENSANUT has a 2016 wave, this was not used since it is a mid-way survey with smaller sample size and some of the questions included in other surveys were not asked. ENSANUT surveys are intended to be undertaken every six years. Nevertheless, given the accelerated increase in the prevalence of overweight and obesity, it was decided to conduct a *mid-term* survey between 2012 and 2018 to monitor the health and nutritional status of the population (National Institute of Public Health, 2016). As a result, the 2016 survey was designed differently from the 2012 and 2018 surveys which do not allow us to make comparisons in IOp measures across three points in time. For example, the 2016 survey did not collect data from two States: Colima and Oaxaca. Instead, additional observations from Chiapas, Tabasco and Veracruz were added to substitute data from Oaxaca (Romero-Martínez et al., 2017). It is not clear how data from Colima were replaced. This affects comparisons when using geographical regions, as well as information at the State level. Additionally, this re-assignment has important implications if information about indigenous population is used.

1.3.1 Key variables

The units of analysis throughout are adults, defined by the survey as those aged 20 to 69 years old at the time of data collection. We use valid data from 34,758 individuals in the 2012 survey and 14,839 from 2018.

Outcomes

BMI and WC are used as proxies of nutrition and adiposity-related health outcomes, although these indicators differ in what they specifically measure. BMI is basically the ratio of weight to height and is the most common measure of overweight and obesity due to its availability and simplicity of measurement. Nevertheless, BMI does not consider the body fat distribution and the mass of abdominal fat (visceral fat), this can over-and under-estimate body fat (Dalton et al., 2003). For example, people with considerable muscle mass will have a higher BMI, whereas people with lower mass, for example, elderly people, will have a lower BMI. To overcome these concerns, we also use individual's WC, which accounts for intra-abdominal fat mass. Both indicators are accurate predictors of diabetes (Vazquez et al., 2007), but WC provides a better approximation to coronary heart disease risk (Flint et al., 2010). Anthropometric measurements were taken by trained and specialised staff from the National Institute of Public Health (INSP) in Mexico. Weight, height, and waist circumference were measured twice, thus we took the average of both measures. Pregnant women, individuals who reported having problems relating to measurement procedures, and individuals with biologically implausible values for BMI (BMI$<$10 and BMI$>$59 (González et al., 2013)) and WC ($<$51cm and $>$190cm (Jacobs et al., 2010)) were excluded from the analysis. This amounts to 493 observations for 2012 and 312 for 2018.

Circumstances

In practice, the *ex-ante* approach focuses only on the total contribution of circumstances while efforts are unobservable factors. The set of circumstances chosen for this study incorporates the normative framework embedded in the Mexican Constitution, where the fundamental principles are established, and the *possibilist criterion* (Ramos et al., 2016), which claims that contextual factors matter and should be taken into account, for example, access to basic public services such as running water, electricity, sanitation, etc. The first article of the Mexican Constitution stipulates that *"any discrimination based on ethnic or origin, gender, age, disability, social or health condition, religion, opinions, sexual preferences, marital status or any other that threatens dignity is prohibited"* (Mexican Constitution, 2017). It is important to add that the definition of circumstances is a matter of debate, as mentioned in the review of the literature by Jusot et al. (2019). In this piece of research, the circumstances to be included are based on the normative and legal grounds stated in the Mexican Constitution, the document that sets out the fundamental principles and social rights all Mexicans are entailed, rather than in a purely statistical sense, thus, it is more in line with the definition of *formal* equality of opportunity(Williams et al., 2000), in the sense that no legal barriers should exist to access equal basic rights. This follows John Roemer's strategy of drawing on the socio-legal context for the analysis to help define where the responsibility cut should be drawn. Thus, circumstances encompass proxies of the right not to be discriminated against based on ethnicity, to have running water in the household and to social protection in health. As well as the parent's health condition, and the characteristics of where individuals live, such as their urbanity, level of deprivation and geography.

Ethnicity is a characteristic that people cannot choose. Indigenous people in Mexico are often treated unequally in social and economic terms, although they are entitled to the same rights as non-indigenous people. In this study, ethnicity is claimed to be an illegitimate cause for observing health inequalities. The ethnic condition was defined according to the National Commission for the Development of Indigenous People of Mexico (Comisión Nacional para el Desarrollo de los Pueblos Indígenas, CDI, using its Spanish acronym), which asserts that indigenous people are those who speak at least one indigenous language. **Health insurance** is said to be an illegitimate cause of inequalities since it is a constitutional right that was first established in the Mexican Constitution in an amendment to Article 4th in 1983 that stated "every person has the right to health protection. The law will define the ways and means for access to health services and will establish the concurrence of the Federation and the federated entities in matters of public health"(Mexican Constitution, 2017). Nevertheless, it was not until the 2003-reform of the country's General Health Law that social protection in health was effectively exercised. This reform explicitly adopted social inclusion, equality of opportunity, individual autonomy, financial justice, and social responsibility as its ethical values (Frenk et al., 2015). We included a categorical variable that indicates whether the individual is affiliated with a public or private health institution or if the person has no health insurance whatsoever[4]. **Parental diabetes and hypertension** is also defined as a circumstance that proxies health conditions inherited from parents and acquired behaviours through exposure. It also reflects *genetic luck* (Dworkin, 1981b; Lefranc et al., 2009). These circumstances are used to reflect the role of familial predisposition to obesity (Nielsen et al., 2015) and therefore account

[4]This has been framed as a circumstance because given that the Mexican health system is fragmented into several health institutions, the quality of care is very heterogeneous and this affects people's health. The Mexican health system is primarily divided into public/private spheres, but within the public system, there are six institutions that provide health and social care. These institutions are the Mexican Social Security Institute (IMSS); the Civil Service Social Security and Services Institute (ISSSTE); Health Ministry programmes, such as *Seguro Popular* or INSABI (the *Seguro Popular* programme was targeted for people with no health insurance and started in 2003. A reform took place in 2019 and the *Seguro Popular* programme disappeared and the Institute of Health for Welfare (INSABI) was created to substitute *Seguro Popular*); the state-owned petroleum company: Mexican Petroleum (PEMEX); the Secretariat of National Defence (SEDENA) and the Secretariat of Navy (SEMAR). Membership in these institutions depends on people's jobs. Examples of affiliations include: people working in the informal sector enrolled in the *Seguro Popular* programme; private company workers affiliated to the IMSS; secondary-level teachers working in a public school are insured by the ISSSTE. Workers of PEMEX, SEDENA or SEMAR receive health and social care in their own institutions. Senior public servants tend to have major private medical insurance (León-Cortés et al., 2019; WHO, 2017b).

for the inherited environment and behaviours present within the household. These are binary variables that indicate whether either the mother or father reported to have been medically diagnosed with diabetes or hypertension.

Despite the fourth article of the Mexican Constitution declaring that *"everyone has the right to access and dispose of clean water for personal and domestic consumption in a sufficient, healthy, acceptable and affordable way"* (Mexican Constitution, 2017), by 2015, 5.6% of the Mexican population declared not to have running water in their households. Furthermore, empirical evidence highlights that the OWOB situation in Mexico is driven, in part, by the high intake of SSBs. We argue that this behaviour could be partially driven by the lack of **availability of running water inside the house** due to this constitutional right not being guaranteed by the government[5]. As such, this household characteristic was included, by itself, as a circumstance. Another circumstance that captures the *geography of opportunity*, a concept that describes how the area and geographical space where people live condition access to opportunities (Rosenbaum, 1995) is included and proxied by the level of **social deprivation at the State level**. This variable is a weighted index that considers access to education, health, basic services, and housing spaces at the State level(Consejo Nacional de Evaluación de la Política Social, 2007). The index is estimated by the National Council for the Evaluation of Social Development Policy (CONEVAL, for its acronym in Spanish) every five years. Thus, we used the 2010 and 2015 indices.

Additionally, we included as a circumstance the **geographical region** where people live. This is so because Mexico is a country with a noticeable North-South divide, with the North being more economically advantaged. Recent studies about inequalities in access to public goods and health found that place of residence matters (Altamirano et al., 2018; Monroy-Gómez-Franco et al., 2020, 2021; Plassot et al., 2022). Thus, we included region together with **urbanity** as potential sources of illegitimate inequality. The 32 Federal States of Mexico were grouped into six regions, see Figure A.0.1in the Appendix, as well as Table A.2.1 that provides further details about the circumstances variables.

All these variables have been specifically chosen and titled *circumstances* because they represent illegitimate sources of disparities. This connotes the idea that the lack of running water or health

[5] A clear case is found in San Cristóbal de las Casas, a town in the South-eastern state of Chiapas in Mexico where families were reported to consume more Coca-Cola than bottled water for hydration (López et al., 2018). This appears to be due to a combination of a lack of water, the liquid is heavily chlorinated and a higher supply of Coca-Cola than bottled water that results in the former being cheaper to purchase than the latter (Pliego, 2019).

insurance should not be influenced by personal choices, the labour market, or the political party governing. The definition of a circumstance used in this paper entails a combination of those given characteristics that people cannot change, but also those factors that should guarantee an equal playing field for everyone before exerting any effort. Therefore the normative and legal framework mentioned when defining each circumstance is of high relevance. For further information about the variables included in the estimation of IOp, see also Table A.2.1 in the Appendix.

1.4 Results

1.4.1 Circumstances and Outcomes

In Table 1.1, we describe the samples in terms of the key variables used in the analysis. For demographics, around 48-42% of the individuals in each sample were men. Also, 66-74% of the people were between 20 to 49 years old, and around 27-34% were older than 50, but younger than 69 years old. Most people in the samples were not from an indigenous ethnicity (93-94%). In terms of social care in health protection, by 2018 at least 16% of the adult population were not affiliated with any public or private health institution. Of those affiliated with a public institution, most of them were subscribed to the IMSS or the former Seguro Popular programme. Only 2% of the adults received private health services in 2018. Regarding familial factors, across the two years, the proportions of individuals with a father or mother, not diabetic or without hypertension decreased. For diabetes, this decline was from 82% to 78% in fathers and 75% to 71% in mothers. For hypertension, the proportions decreased from 82% to 74% for fathers and 67% to 57% for mothers. In terms of household conditions, the proportion of individuals that have piped water inside of their household increased from 69 to 75%, conversely individuals with piped water outside of their household or no piped water decreased (24% to 20%, and 7% to 5%, respectively). With regards to State deprivation, around half sample lived in States considered to be of low and very low deprivation. Finally, most of the adults lived in urban, metropolitan areas or in the Central Region.

Figure 1.1 and Table 1.2 display the distribution of the outcomes by sex and survey year. Both the average BMI and WC across years have increased for women and men. Average BMI rose from 29 to 29.6kg/m^2 in women and from 27.8 to 28.3kg/m^2 in men. The mean WC for women was 92.6cm in 2012 and 95.6 in 2018. For men, WC increased from 94.8 cm in 2012 to 97.8 cm in 2018. Women were observed to have a higher BMI than men. All these average values were above the cut-off points for normal weight, reflecting the presence of overweight and obesity. A healthy BMI is between 18.5-24.9 kg/m^2, while excess weight is identified when BMI is between 25-29.9 kg/m^2,

Table 1.1: Descriptive statistics of the circumstances by health outcome and year

	BMI 2012	**BMI 2018**	**WC 2012**	**WC 2018**
Men	0.48	0.42	0.48	0.43
Age groups				
20 to 29	0.28	0.21	0.27	0.21
30 to 39	0.25	0.21	0.25	0.21
40 to 49	0.21	0.24	0.22	0.24
50 to 59	0.16	0.19	0.17	0.20
60 to 69	0.10	0.14	0.10	0.14
Ethnicity				
Non indigenous	0.93	0.94	0.94	0.94
Health Affiliation				
None	0.25	0.16	0.25	0.17
IMSS	0.32	0.34	0.33	0.34
ISSSTE	0.06	0.07	0.06	0.07
Seg. Pop.	0.35	0.40	0.35	0.39
PDM	0.01	0.01	0.01	0.01
Private	0.01	0.02	0.01	0.02
Parents' health				
Father non diabetic	0.82	0.78	0.82	0.78
Father without hypertension	0.82	0.74	0.82	0.74
Mother non diabetic	0.75	0.71	0.75	0.71
Mother without hypertension	0.67	0.57	0.67	0.57
Water availability				
Piped inside household	0.69	0.75	0.69	0.75
Piped outside household	0.24	0.20	0.24	0.20
No piped water	0.07	0.05	0.07	0.05
State Depriv.				
Very high State deprivation	0.10	0.10	0.10	0.10
High State deprivation	0.25	0.22	0.25	0.22
Medium State deprivation	0.16	0.17	0.16	0.17
Low State deprivation	0.31	0.33	0.31	0.34
Very low State deprivation	0.19	0.17	0.19	0.17
Geo. Region				
Urban-Metrop.	0.79	0.79	0.79	0.79
Northwest	0.08	0.09	0.08	0.09
Northeast	0.18	0.17	0.18	0.18
West	0.19	0.19	0.19	0.19
Centre	0.33	0.32	0.33	0.32
South	0.17	0.17	0.16	0.16
Southeast	0.06	0.06	0.06	0.06
N	34,265	14,517	33,353	14,246

Notes: N=Number of observations
All are binary variables. Depriv=Deprivation. Geo=Geographical
Seg. Pop= Seguro Popular. PDM=Pemex, Defensa and Marina

and obesity when BMI>30 kg/m^2 (WHO, 1995). For WC, excess adiposity is identified when the WC is above 90 cm in men, and above 80 cm in women (Alberti et al., 2009).

Table 1.2: Descriptive statistics of the health outcomes split by sex and year

	2012		**2018**	
	BMI	WC	BMI	WC
Women				
N	19,906	19,024	8,282	8,006
Min	13.3	52.3	15.0	51.2
Mean	28.9	92.4	29.6	95.7
Max	58.9	168.6	58.9	164.6
SD	5.8	13.4	5.9	13.9
Men				
N	14,359	14,329	6,235	6,240
Min	11.1	53.0	15.7	53.7
Mean	27.7	94.7	28.3	97.8
Max	57.4	190.0	56.1	187.4
SD	4.9	12.9	4.9	13.3

Note: N=number of observations
BMI=Body Mass Index. WC=Waist Circumference
Min=Minimum value. Max=Maximum values. SD=Standard deviation

1.4.2 Regression models

Results from modelling BMI and WC conditional on circumstances depicted in Equation (1.6) show that nutrition-related outcomes were greater for non-indigenous people, with a higher magnitude in WC. In terms of health care protection, there were differences across the years. For example, compared with people that were not affiliated with any health institution, adults affiliated with public institutions had a lower BMI or WC in 2012. Nevertheless, in 2018 only those affiliated with the ISSSTE had a lower BMI or WC. The regression models also showed a systematic and statistically significant negative relationship between parents' diabetic condition and health outcomes, with a higher magnitude for WC. There was also a negative relationship between our outcomes and parents' hypertension status across both years, but not all these relationships were statistically significant. Individuals that had no piped water available in their households had lower BMI or WC than people that had piped water available in their homes. Although there were some differences across categories, in general, the lower the level of deprivation, the higher the BMI or WC, with higher magnitudes in WC. We also found that living in urban or metropolitan areas was positively correlated with BMI or WC. Finally, when compared with the Northwest region, those individuals living in the Northeast, West, and Central regions had a lower BMI or WC across both years. However, adults that lived in Southeast regions had a higher BMI than those living in the Northwest in 2018. It is worth noticing that in 2018, individuals living in the South and Southeast region had a higher and lower WC, respectively, than those living in the Northwest region. Results for these models can be found in the Appendix, Tables A.2.1-A.2.4.

1.4.3 Inequality of opportunity in BMI and WC among Mexican adults

Table 1.3 shows the levels of *ex-ante* inequality of opportunity in BMI and WC for Mexican adults in 2012 and 2018. Absolute and relative levels are displayed with the latter expressed as a percentage. Overall, relative IOp in both outcomes has remained stable across time, at around 3.5% for WC and between 4.0% and 4.3% for BMI. Absolute inequality did not change across time in both outcomes, but total inequality decreased in BMI.

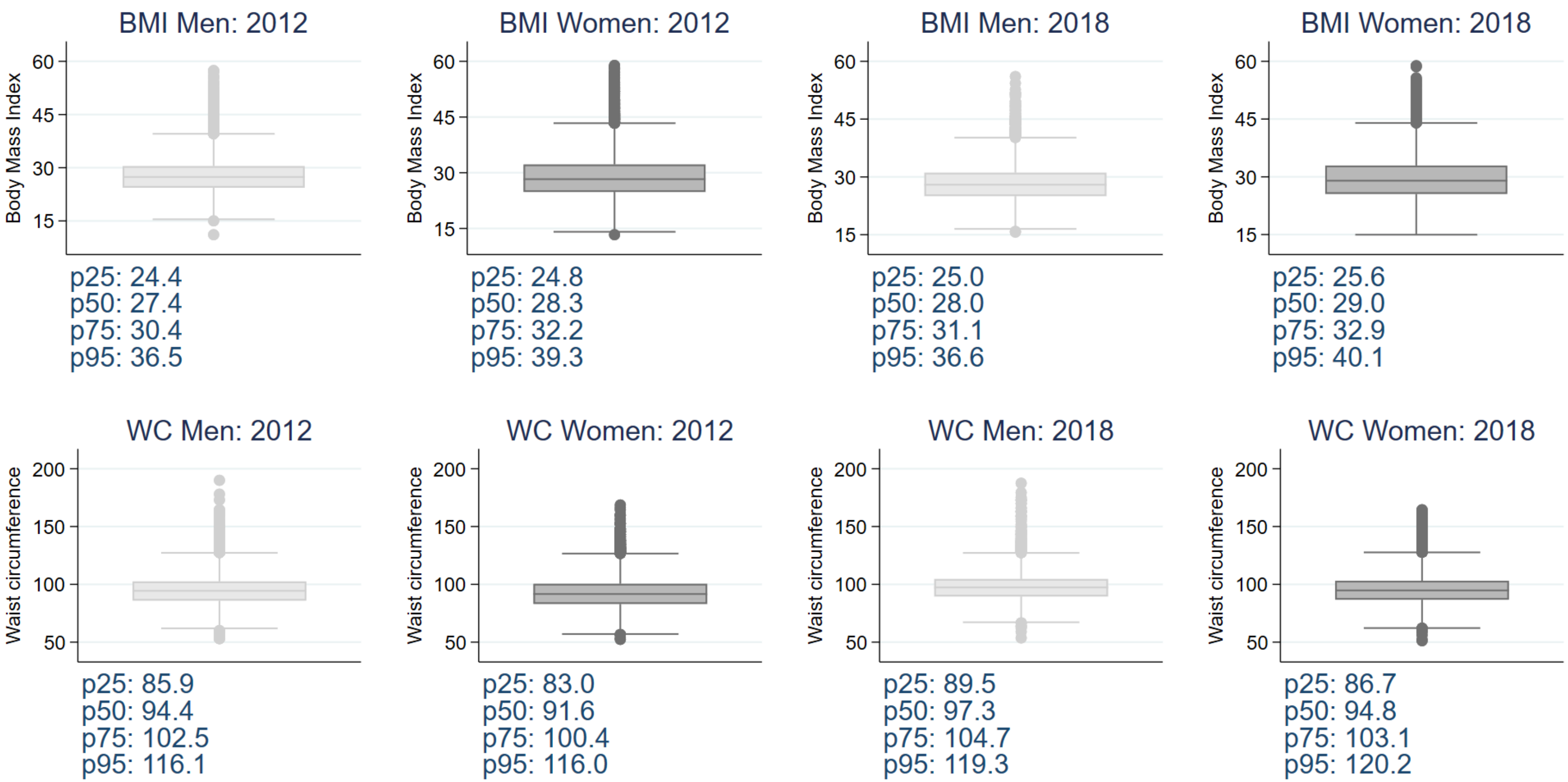
Table 1.3: Absolute, Total and Relative Inequality of Opportunity

Outcome-Year	***Absolute***	*BSE*	***Total***	*BSE*	***Relative (%)***	***Observations***
BMI 2012	0.0007***	0.0000	0.0174***	0.0001	4.022	27,612
BMI 2018	0.0007***	0.0000	0.0171***	0.0002	4.093	12,644
WC 2012	0.0003***	0.0000	0.0095***	0.0001	3.157	26,808
WC 2018	0.0003***	0.0000	0.0095***	0.0001	3.157	12,392

Notes: BSE=bootstrapped standard errors (500 replications)

+ $p<0.1$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$

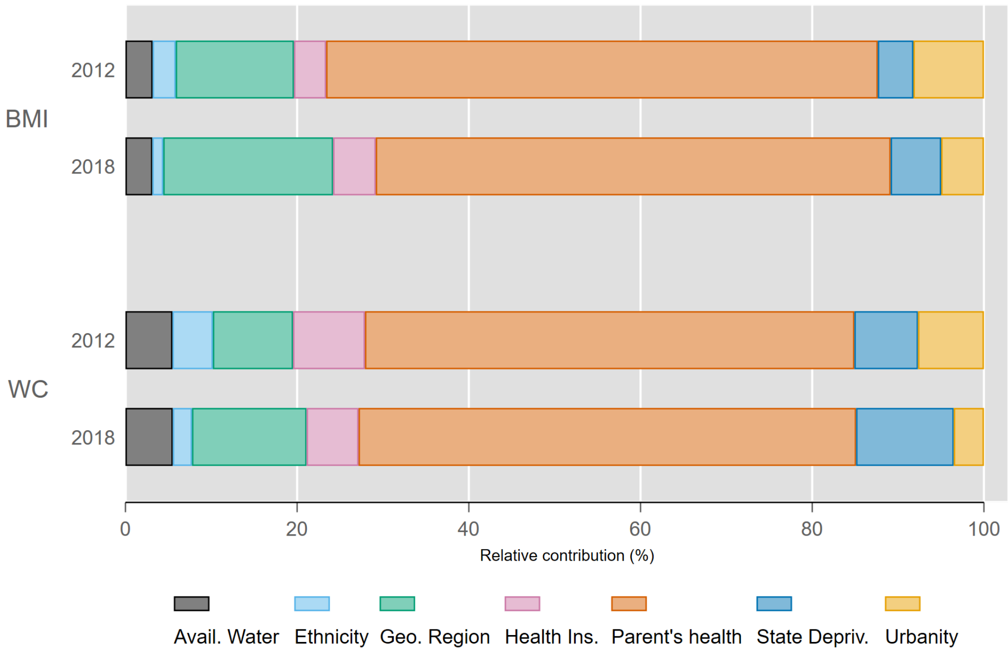
Figure 1.1: Distribution of BMI and WC split by sex and year



1.4.4 IOp drivers

Figure 1.2 shows the results of using the Shapley-Shorrocks decomposition method to identify circumstances that contribute the most to *ex-ante* IOp. Across survey years and outcomes, parent's health was the circumstance that accounted for most of *ex-ante* inequality, at around 57-64%. The second most relevant circumstance was the geographical region where people lived (9-13%). For both outcomes, the relevance of state deprivation, urbanity, and social protection in health and ethnicity decreased across the years. Overall, the relevance of water availability remained unchanged over time, but it is of higher importance for WC than BMI. Specific relative contributions can be found in Table A.3.1 in the Appendix.

Figure 1.2: Relative contribution of each circumstance to IOp split by outcome and year



1.4.5 Distributional analysis: going beyond the mean

When applying Equation (1.12) to RIF models, we found similar patterns in our regression models. Non-indigenous people had higher outcomes than indigenous people, individuals whose parents were not diagnosed with diabetes or hypertension had lower BMI and WC values. Although there were some relevant differences at the bottom and top percentiles. For example, compared with individuals with no health protection, individuals at the p25th had a higher BMI or WC if they were affiliated with a public health institution, except for those affiliated at PEMEX, MARINA and DEFENSA. This result was expected since better health outcomes are observed in people receiving health care in these institutions (WHO, 2017b). Another relevant difference was that at the mean, those individuals with no piped water had a lower BMI compared with those adults that had piped water available inside of their household. But this relationship was only positive for the p95th of the BMI distribution. Overall, these circumstances explained more of the variation of the outcome in the lower percentiles: only 3-4% of the BMI variation was explained by this set of variables at the p95th, while 10% of the total variation was explained at the p25th. Results for all models are found in the Appendix, Tables A.2.2 to A.2.5.

Table 1.4 shows the level of IOp across the distribution of each outcome, for both survey years. Overall, absolute inequality increased across both outcomes' distributions. Thus, by relaxing the inequality neutrality assumption, we found that circumstances mattered more for individuals at the top than at the bottom of the distribution. Total inequality was highest at the bottom and upper levels of the distribution for both outcomes and years. The highest relative inequality associated with circumstances occurred at between the 50th and 75th percentiles. Except for IOp in BMI at the p90th in 2018, which was the highest. This is so because both absolute and total inequality were the highest for that outcome and year. Comparing across years we found that, while absolute inequality slightly changed across the WC distribution, absolute inequality in BMI increased the most at the 95th percentile. Total inequality did not change much from 2012 to 2018 across both distributions.

When identifying the relative contribution of each circumstance towards *ex-ante* IOp across the whole distribution, it was observed that parents' health conditions were, once more, the main driver of illegitimate disparities. Although there were differences in magnitude across percentiles and years. For example, in BMI and WC for 2012, Figure (1.3) shows that the contribution of parents' health was at around 48-61%. However, there were changes in 2018 so that their relevance was higher at the lower parts of both outcomes' distribution (25th percentile), at around 60-66%. The second driver was the geographical region where people lived, again with differences across outcomes and survey years. Comparing across the distribution, the relevance of geography was higher at the upper parts of the distribution (e.g., the 75th and 95th percentiles), around 17% for BMI in 2012 and 24-40% in 2018, and 14-17% in WC for 2012 and 26-30% in 2018.

The relevance of health insurance was particularly high in the 25th percentile, but only in 2012 for both outcomes (15% and 22% for BMI and WC, respectively). The analysis showed that State deprivation was of higher importance for *ex-ante* IOp in WC than in BMI, and of relevance at the middle of the distribution. Water availability was a circumstance that mattered more at the

Table 1.4: Inequality of opportunity across different percentiles

Outcome-Year	***Absolute***	*BSE*	***Total***	*BSE*	***Relative (%)***	***Obs.***
BMI q25 2012	0.0006***	0.0000	0.0408***	0.0004	1.56	27,612
BMI q50 2012	0.0007***	0.0000	0.0249***	0.0000	2.63	27,612
BMI q75 2012	0.0008***	0.0000	0.0266***	0.0002	2.83	27,612
BMI q95 2012	0.0011***	0.0000	0.0475***	0.0011	2.41	27,612
WC q25 2012	0.0003***	0.0000	0.0262***	0.0002	1.28	26,808
WC q50 2012	0.0004***	0.0000	0.0152***	0.0000	2.57	26,808
WC q75 2012	0.0004***	0.0000	0.0146***	0.0001	2.87	26,808
WC q95 2012	0.0005***	0.0000	0.0263***	0.0006	1.89	26,808
BMI q25 2018	0.0007***	0.0000	0.0400***	0.0005	1.70	12,644
BMI q50 2018	0.0007***	0.0000	0.0229***	0.0000	3.10	12,644
BMI q75 2018	0.0009***	0.0000	0.0277***	0.0002	3.11	12,644
BMI q95 2018	0.0016***	0.0000	0.0473***	0.0016	3.40	12,644
WC q25 2018	0.0004***	0.0000	0.0226***	0.0003	1.91	12,392
WC q50 2018	0.0003***	0.0000	0.0122***	0.0000	2.51	12,392
WC q75 2018	0.0004***	0.0000	0.0150***	0.0001	2.76	12,392
WC q95 2018	0.0007***	0.0000	0.0300***	0.0011	2.35	12,392

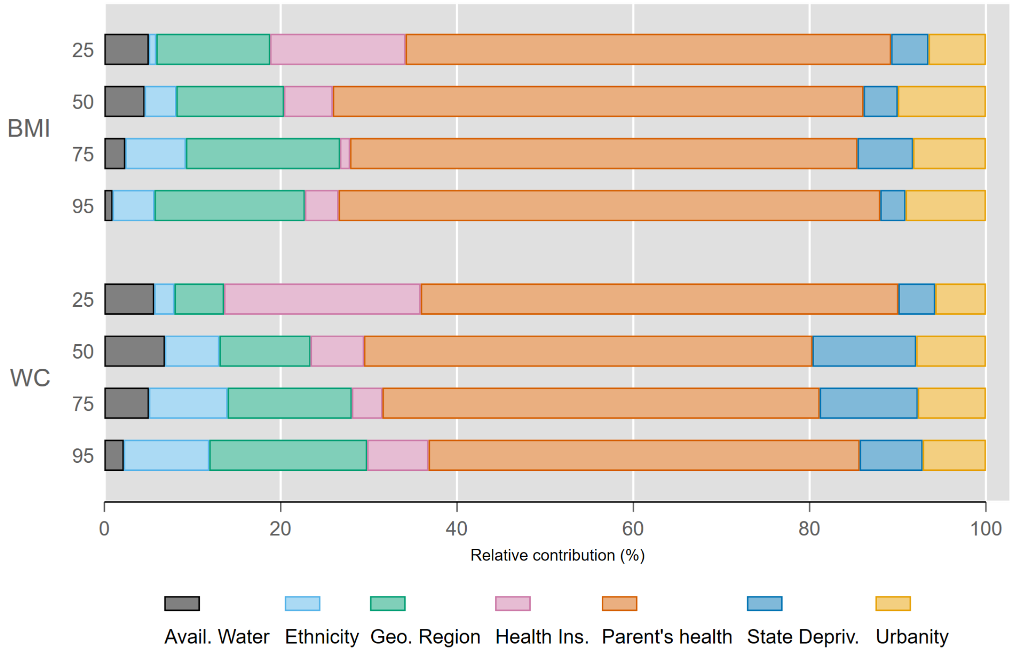
Notes: BSE=bootstrapped standard errors (500 replications). Obs=Observations
+ $p<0.1$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$

upper parts of the BMI distribution. Although, the relevance of this variable at the 25^{th} and 95^{th} percentiles for WC in 2018 was similar. The importance of urbanity for both outcomes remained constant across their distribution in 2012, and some differences across percentiles in 2018. The importance of ethnicity was higher at the upper parts of both outcomes' distribution. To be indigenous or not mattered little to IOp at the 25^{th} percentile of both outcomes and survey years. Specific relative contributions at different parts of both distributions can be found in Table A.3.2 in the Appendix.

1.5 Discussion

Following the canonical work of John Roemer, subsequent research has acknowledged that not all inequalities are unfair. A first step in identifying illegitimate IOp is to disentangle the extent to which inequalities in outcomes are due to circumstances. In this regard, if circumstances play a role in achieving a certain outcome, individuals face unequal *playing fields*. This analysis has measured the level of inequality related to circumstances in two nutrition-related health outcomes among the Mexican population, one that has typically looked at excess weight and another at excess adiposity. To measure IOp, we first assumed within-type transfer insensitivity (Ferreira et al., 2011), which implies that variation in outcomes within individuals that share the same circumstances does not

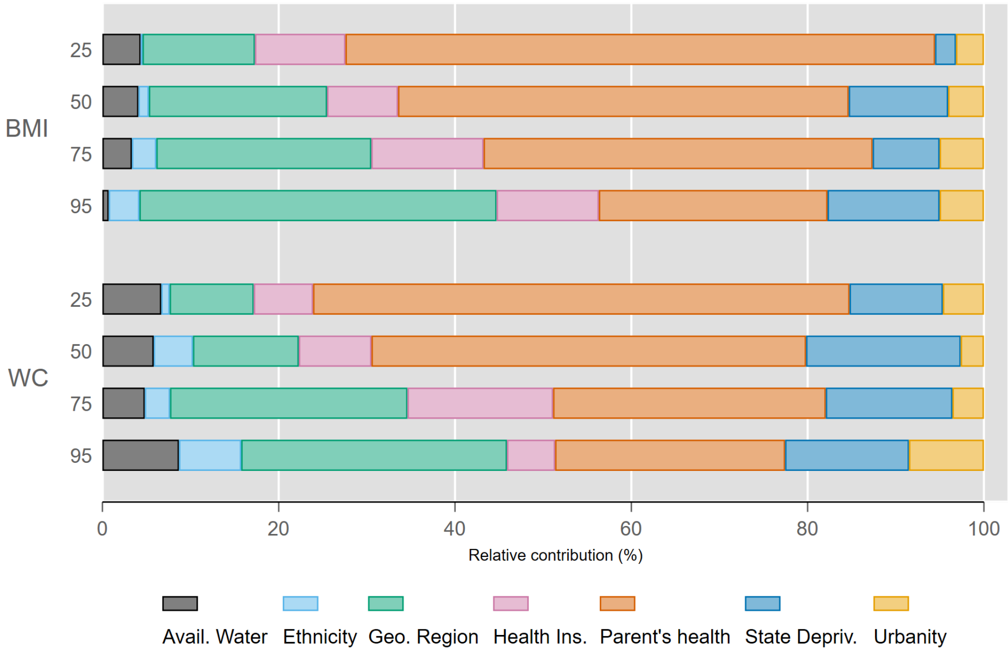
Figure 1.3: Relative contribution of each circumstance to IOp at different percentiles in 2012



matter. This is a sensible assumption to explore disparities in outcomes across people with different opportunity sets.

Overall, the level of relative inequality attributed to circumstances for BMI and WC is similar and around 3-4%. These estimates only considered a normative-based opportunity set. Variables were selected upon fundamental rights included in laws, acts and most importantly, the Mexican Constitution. Using these normative-related variables and acknowledging the potential omission of other relevant circumstances, these estimates represent lower-bound levels of IOp (Ferreira et al., 2011). Our estimates show the total effect of circumstances on BMI and WC, reflecting both, the direct effect of circumstances and their indirect effect via efforts. The results indicate that the direct effect of effort and luck is likely to explain most of the disparities in BMI and WC. Since the *ex-ante* approach conceptually takes all efforts into account (both observed and unobserved), there might still be unobserved circumstances. β_1, in the reduced form shown in Equation (1.6), estimates both the direct effect of circumstances and their indirect effect through efforts. Since there might be omitted circumstances, a lower bound of ex-ante IOp is estimated. One circumstance that could be of high relevance, but that was not included in this analysis due to the lack of data, is information about epigenetics, how parents' behaviours and living conditions affected the adiposity-related genes inherited by their children. Also of relevance might be adverse childhood

Figure 1.4: Relative contribution of each circumstance to IOp at different percentiles in 2018



conditions, such as the presence of food insecurity within the household, or the role of economic shocks and their effect on maintaining good nutrition. This is important since the adults analysed in this study were exposed to the 1988 economic crisis and the market-oriented policies that followed suit. There is evidence that this affected severely the quantity and quality of the food consumed by families and individuals. These are potential circumstances that were not included but might be relevant for the estimation of IOp in nutrition-related outcomes. Furthermore, it has to be noted that this is a cross-sectional analysis that focused only on adult individuals (20 to 69 years old) and it is very likely that IOp in these outcomes varies across different stages of the lifespan. This aspect is missing in this analysis. Early-life circumstances shape later-life opportunities and outcomes (Aizawa, 2020) and efforts are mediating efforts in this trajectory (Jusot et al., 2013). These two effects are not explicit here. Furthermore, while other studies conceptualised sex and age as illegitimate sources of inequalities and obtained higher levels of IOp (Davillas et al., 2020a), we defined them as control variables to capture biological determinants of health status, following Jusot et al. (2013), and included them as fixed terms in our models.

Although *ex-ante* inequalities proved to be relatively low in this analysis, we argue that by no means this implies that it is justifiable not to eliminate them. There is room for policy prescription, for example, one related to inequality reduction. Applications of the social theory framework to income-related inequality predict that if utility and wealth show an increasing and concave function, marginal utility diminishes with respect to wealth. This means that an extra pound given to a rich person increases utility less than the same extra pound given to a poor person. This principle applied to health means that giving an additional amount of health, e.g., a QALY (quality-adjusted life year), to a person whose QALYs are 75 years is less valuable than giving a QALY to someone with an expectancy of only 30 years. Hence, if health shows a concave relationship, the Pigou-Dalton (PD) or *between-type transfer* principle holds. This is a desirable property of inequality measures and requires absolute IOp, θ_a, to rise if a transfer of health from someone ill to someone healthy is made (Ferreira et al., 2011). Thus, there is room for social planners to design policies that increase social welfare via transferring health from healthy to ill individuals and still preserving the average health of the population. Notwithstanding, the relationship between our health outcomes and social welfare is not concave along all its distribution. Social welfare will certainly increase if the BMI of an underweight adult rises. The clinical literature acknowledges that if the BMI of an underweight adult (BMI$<18.5kg/m^2$) increases, those individuals will reach a normal and thus healthy weight. However, once BMI passes the cut-off of $25kg/m^2$, individuals are not healthy, as they are overweight, and become unhealthier, where BMI goes beyond 30-$35kg/m^2$ (e.g., obese or extremely obese). This is also the case for WC, which also shows a nonlinear relationship with respect to social welfare. Social utility increases as WC increases but up to a threshold of 90 cm for men and 80 cm for women. The PD principle does not hold when the mean-based approach is followed, given the nature of our outcomes. It is worth noting as well the assumption of inequality neutrality or *within-type transfer insensitivity* (Ferreira et al., 2011) in this approach, which implies that disparities in outcomes among people that share the same circumstances is irrelevant. Ferreira et al. (2011) state that the absolute inequality measure is deliberately insensitive to within-group inequality and that this is the focus axiom for *ex-ante* IOp. This is consistent with a linear utility function in which each person is given equal weight in the social welfare function. Therefore, any policy focused on health transfers among individuals that are similar in their circumstances will have no effect on increasing social welfare.

We tackled the increasing and decreasing nature of utility with respect to health concerns in two ways. First, we undertook a distributional analysis, where IOp is measured at several points of the BMI and WC distribution, thus allowing inequality aversion. The distributional analysis allowed us to examine the different weights that circumstances have across the BMI and WC distribution. Thus, we have assessed IOp for those individuals at the bottom of the distribution, presumably underweight or with lack-of-adiposity, and at the top of the distribution, where individuals suffer overweight, obesity or excess adiposity. Second, to overcome the non-linear nature of our outcomes, we dichotomised the outcomes according to the clinical cut-off points for overweight and excess adiposity according to an individual's BMI and WC and measured IOp following Paes de Barros et al. (2008). Thus, we estimated the dissimilarity index (D-Index) using a logit function and including the same set of normative-based circumstances as well as sex and age fixed effects. This absolute inequality measure is ideal for binary outcomes. In this way, the focus was only on over-nutritional outcomes. The dissimilarity index indicates the fraction of opportunities that need to be reallocated from the better-off to the worse-off groups to achieve equality of opportunity holding constant the population average outcome. Hence, it measures the opportunity set gap. This index has desirable properties, for example, it is sensitive to redistribution of health from non-vulnerable to vulnerable groups, but it is insensitive to redistribution of health within people that share the same circumstances (Bruoni, 2016). It is also scale-invariant, exhibits anonymity and is invariant to population replication. For example, the results of this analysis, shown in Table A.4.1 in the Appendix, indicate that 7% (5.6%) of opportunities were needed to be transferred in 2012 (2018) from the best-off to the worse-off groups to observe equality of circumstances in BMI.

A third potential option that could be explored in further research is to estimate inequality using polarisation[6] (Apouey, 2007) or entropy measures (Contoyannis et al., 2007). One could split the BMI distribution into different ordinal categories, for example, low-weight, normal weight, overweight and excess weight, and estimate the degree of polarisation in the distribution of the new ordinal variable. Although this is a potential solution to the non-linear nature of the outcomes and the difficulties that mean-based inequality measures face[7], how to incorporate the theoretical IOp framework to define reference groups for comparison has not been studied. We also found that illegitimate inequality is driven by people's parental conditions and determined by the geographical place where they live. Particularly, having parents that have been diagnosed with diabetes greatly contributes to higher IOp. This could potentially be associated with mechanisms in which parents with obesity-related diabetes pass to their children certain physical characteristics that lead to inter-generational obesity (Brisbois et al., 2012; Wrotniak et al., 2004) and it can also be linked to evidence about familial predisposition to obesity (Nielsen et al., 2015; Teran-Garcia et al., 2013). In Mexico, the prevalence of type two diabetes in adults is around 13-22% (Meza et al., 2015; Saeedi et al., 2019) and it has been estimated that 90% of the cases are linked to OWOB

[6]Polarisation measures are ideal to describe situations where the extremes of a distribution grow, and the middle shrinks. While inequality measures look at variations across distributions.

[7]One of the caveats of *some* inequality measures is that these are appropriate for cardinal data and are based on the mean of the distribution. Thus, relevant distributional aspects are missed when measuring disparities using ordinal variables.

(Dávila-Torres et al., 2015; Health, 2010), which suggests that familial context matters. A recent study that compared growth trajectories and children's caloric intake according to post-partum mother's BMI found that social environmental factors like food landscape might play a decisive role in shaping children's obesity (Téllez-Rojo et al., 2019). This evidence inter-generational transmission of obesity-prone behaviours. Given the implicit egalitarian principle behind the *ex-ante* approach, compensatory policies should therefore exist to dampen the effect of unequal early-life circumstances. These could take the form of differentiated healthcare policies or interventions that focused on obesogenic environments in households during pregnancy and childhood early life stages (Haire-Joshu et al., 2016).

The geographical region where individuals live is the second main driver of disparities. Where people develop their life matters and it is more relevant for people in the upper parts, 75^{th} and 95^{th} percentiles, of both BMI and WC distributions. This sheds light on geographical differences in risk exposures to worse health that might potentially be attenuated by localised and differentiated health and social policies. This complements and coincides with recent findings about the contribution of geography *per se* on excess adiposity in England. Davillas et al. (2020b) explored the relative contribution of geographic areas on excess adiposity and found that the role of geography is more pronounced and relevant for individuals at the top of the BMI and WC distribution. This highlights that factors beyond individual control can be modifiable via health, social and economic local programmes, and interventions.

Water availability was seen to be relevant for inequalities at the higher parts of both, the BMI and WC distribution. Given the evidence that exists about the connection between no running water availability and excessive drinking of sweetened beverages (López et al., 2018; Pliego, 2019), we hypothesised that the limited extent of water access, as a fundamental right, could be related to higher inequalities. In Mexico, everyone is entailed access to running water, regardless of their economic position or location of residence. Nevertheless, only 75% of the national population had pipped water inside of their household in 2018, and still, 5% had no piped water either inside or outside of their households. This lack of water availability contributed to explaining around 10% of illegitimate inequalities across the distribution of both outcomes, particularly in WC at the bottom and top end of its distribution. In addition, despite the entitlement of the right to protection of health, this analysis found that unfair inequalities in BMI and WC are boosted by the lack of social protection in health. Health insurance explains around 7-20% of *ex-ante* inequalities. Its relevance was higher at the bottom of the distribution in both outcomes. This could potentially be linked with evidence that points out that the lack of primary healthcare, which is aimed to enhance health promotion and timely detect obesity and overweight, is related with higher rates of diabetes and hypertension in Mexico (Alcalde-Rabanal et al., 2018).

This analysis is not without limitations. One important drawback is the data. ENSANUT neither collected retrospective data about the familial background nor collected data in a panel format. More robust influence could be obtained with panel data or retrospective information about parents' health backgrounds. In addition, parental diabetes and hypertension condition is self-reported, which might potentially induce bias through measurement error. Nevertheless, the proportion of parents with diabetes (18-26%) and hypertension (24-43%) in this study are relatively similar to the expected national prevalence rates of 13–22% for diabetes (Meza et al., 2015; Saeedi et al., 2019), and 13% to 44% for hypertension (Sudharsanan et al., 2019). Overall, reflecting on the fact that health outcomes depend on circumstances, efforts, and luck, and that with this analysis we have isolated the role of normative-based circumstances, these findings suggest that efforts and luck might play a predominant role in BMI and WC, although further analyses are needed to confirm this.

In democratic societies, such as in Mexico, equality of opportunities in health is not only desirable but also paramount for social well-being and development. Unequal health outcomes across individuals are not necessarily unfair. Based on ethical grounds, there is a problem if health outcomes depend on people's ethnicity, parental background or unequal access to fundamental rights and services. Within this context, this study has explored another aspect of the acute OWOB situation in Mexico. This analysis has further implications for the economic approach to obesity, which has mostly been studied under a consumer's behaviour view. Obesity has been framed as the outcome of powerful social and cultural forces that promote an energy-rich diet and a sedentary lifestyle (Obesity Institute of Medicine, 1995) or as a side-effect of technology changes or increased female participation in the labour market (Rashad et al., 2004). Although this might be the case, the economics of obesity should also incorporate the social, political, and institutional structures in which people develop their lives and the potential role that governments have, not only in implementing policies to outweigh the downsides of female labour participation or enact changes to health but also in guaranteeing fundamental rights to all citizens. This study claims that unequal opportunities condition further choices and lifestyle decisions. In this regard, further interventions should acknowledge that equalising the playing field is a premise for effective public policies to tackle the OWOB crisis.

Chapter 2

Accumulation and transmission of inequality of opportunity in the double burden of malnutrition: the case of Mexico

Abstract: Using a life-course perspective and based on Roemer's inequality of opportunity framework, the hypothesis of an accumulation and intergenerational transmission of *ex-ante* and *ex-post* inequality of opportunity in malnutrition is tested. This paper measures the evolution of inequalities considering the socioeconomic changes and the evolution of circumstances and efforts experienced by people born between 1983 and 1988 in Mexico. Using a combination of matching and re-weighting methods, a pseudo-birth-cohort is constructed and the effect of circumstances and efforts on inequality of opportunity is disentangled and measured across nutrition-related health outcomes. Results indicate that inequality of opportunity in malnutrition has been a persistent issue across the life course of the birth cohort and that lack of opportunities has been transmitted from parents to children. When disentangling the contribution of circumstances and efforts to inequality in malnutrition, we find that, on average, people's circumstances explain most of the explained variation (72-76%), whereas efforts account for little of the variation (28-24%). We find that circumstances are the main driver of inequality in undernutrition and no consistent evidence that efforts play a dominant role in explaining variation in outcomes associated with overnutrition. The empirical results are relevant for a better understanding of the "economics of obesity".

Keywords: Double burden of malnutrition; Inequality of Opportunity; Matching and re-weighting; Mexico

2.1 Introduction

Over the past few decades, several upper and middle-income countries, Mexico included, have experienced significant epidemiological changes. Concurrently, eating patterns, nutritional status and the disease burden of the population have been radically modified. Somewhat paradoxically perhaps, at the same time, obesity, stunting and anaemia have been observed in populations, households and individuals (Kroker-Lobos et al., 2014; Shrimpton et al., 2012; WHO, 2017a). The determinants behind the coexistence of stunting and obesity in upper and middle-income countries have already been studied. Rapid urbanisation, demographic changes, the modification of dietary patterns and lifestyles are factors closely related to the *nutritional transition* and the *double burden of malnutrition* (Batal et al., 2018; Doak et al., 2005; Popkin, 2001, 2015; Popkin et al., 2012; WHO, 2017a). The nutritional transition occurs when rapid modification of traditional diets and physical activity patterns takes place, usually across socioeconomic and demographic groups. One characteristic is that local traditional eating patterns change towards westernised diets, which are high in fat, salt, and sugar and with low nutritional value. This transition has preceded the double burden of malnutrition so that now stunting and obesity jointly can be observed in the same households, populations, or individuals (Tzioumis et al., 2014).

Political, macroeconomic, and social changes also shape people's health. Mexico is a clear example. The macroeconomic shocks that occurred during the 1980s contributed to the nutritional profile of the population. In 1988, the highest-ever level of inflation (4,030%) was registered. Consequently, purchasing power plummeted by 70%. The stagflation crisis led to the adoption of market-oriented economic policies, including trade liberalisation. In 1994 Mexico subscribed to the North American Free Trade Agreement (NAFTA) with the United States of America (USA) and Canada. This agreement aimed to remove barriers to free trade by eliminating any kind of tariffs on imports and exports between the three countries. Evidence suggests that NAFTA transformed Mexico's food system (Clark et al., 2012). The flow of corn, soybeans, livestock, meat, and feed grains, as well as sugar and sweeteners from the USA to Mexico, increased dramatically. NAFTA has directly (and indirectly) changed Mexico's food supply chain. American direct investment in Mexico also grew, particularly the number of fast-food companies substantially increased. Thus, the Mexican diet changed from a traditional plant-based to animal energy-dense and processed food diet (Clark et al., 2012). A study, that evaluated the effect of food trade between Mexico and the USA on obesity in Mexico, found that exposure to food imports from the USA explained up to 20% of the rise in obesity prevalence among women between 1988 and 2012 (Giuntella et al., 2020). Another study that focused on characterising the effects of the 90s economic crisis on calorie intake in Mexican households found that, in general, the total calorie intake did not change, although the consumption of expensive calories (meat, eggs, milk and soft drinks) increased and inexpensive staples (cereals, legumes, sugars) decreased. The study concluded that high energy and non-nourishing calorie consumption had emerged (Arroyo et al., 2004).

In terms of the social and health conditions, by the end of the 1980s, the Mexican population was facing a high risk of stunting, predominantly in indigenous populations, rural municipalities, in the South and Central regions, and in households with poor conditions and where mothers had a low educational background (Rivera-Dommarco et al., 1995). A study describing the level of iron deficiency among women of reproductive age found that the prevalence of anaemia among women was higher in pregnant compared with non-pregnant women. Results from this research also showed that anaemia prevalence was higher among indigenous women and women living in urban areas (Martinez et al., 1995). Another cross-sectional analysis of feeding patterns of infants in Mexico found that in 1988 the hazard rate for terminating breastfeeding increased by 38% for each increment in the household's category of living conditions at the national level (Long-Dunlap et al., 1995). In terms of health coverage, by 1995 only half of the population had access to health-care coverage (Leal et al., 2002).

The study of the evolution of malnutrition considering the socioeconomic changes and the evolution of opportunities experienced in people born between 1983 and 1988 in Mexico is of high relevance. Not only because, to the best of our knowledge, there are no studies that have focused on analysing i) malnutrition, as a spectrum that includes both under and over-nutrition, ii) the accumulation of socioeconomic-related health inequalities during a life course of 30 years and, iii) the potential transmission of inequalities across generations. But also, because studying the potential accumulation and transmission of health inequalities raises important questions, from a philosophical and practical perspective. The study about IOp and malnutrition has been growing in the last two decades. For example, Aizawa (2019) applied an *ex-ante* approach to IOp to study differences in child malnutrition in ten Asian countries and relying on cluster analysis to partition children according to their circumstances. This analysis found that the magnitude of inequities ranged between 21.7% and 5.9%, being Pakistan the country with the highest level of IOp and the Maldives with the lowest. Further identification of the main drives behind these illegitimate inequalities showed that housing conditions were the most relevant circumstance. Another study about IOp in malnutrition in children under five was carried out by Sanoussi et al. (2020), which studied the case of Congo, Guinea Bissau, and Mali. The analysis found high disparities between the most advantaged and least advantaged groups and that household welfare was the main source of unfair inequalities. Another study about IOp among Egyptian children during the 2000s decade found that although overall IOp decreased, inequalities in nutrition indicators were still a matter of concern (Ersado et al., 2014). Research about *ex-ante* IOp in z-scores of height-for-age (HAZ) and weight-for-height (WHZ) in children under five years of age in Arab Countries and Turkey found that boys and girls kept facing unequal opportunities and that these disparities tended to accumulate. Most of these inequalities were due to parental wealth and education circumstances (Assaad et al., 2012). Recently, Liu et al. (2022) also applied an *ex-ante* approach to IOp in HAZ, weight-for-age (WAZ), WHZ and body mass index z-scores (BMIZ) and found that IOp was of a magnitude of 11.49%, 3.64%, 7.92%, 6.9% and 9.13%, respectively. They also found that family ground and the geographical region where children lived were the main drivers of inequities. Although these studies have studied the case of low and middle-income countries (LMIC) and coincided with finding that circumstances play a relevant role behind disparities, the analyses have focused on examining the early stages of life, mostly on children under five years of age. This might be explained by the unavailability of longitudinal data. The lack of panel data is a common issue in these countries, where nutritional transitions mostly take place. Although in Mexico there is the panel survey: "Mexican Family Life Survey" (MxFLS), its time horizon covers only 10 years, from 2002 to 2012. Hence, the potential life span to be studied is very short and does not allow an analysis of the effect of the 80s and 90s economic policies on individuals' health.

This study aims to overcome the limitations previously described. We use nationally representative surveys that span a longer time period, from 1988 to 2018 and propose an empirical strategy that relies on the use of matching and re-weighting methods to construct a pseudo birth-cohort panel. This strategy allows us to analyse the potential accumulation and transmission of inequality of opportunity (IOp) in malnutrition-related health outcomes over a period of 30 years. The use of these surveys is relevant since it allows us to exploit the rich data about food consumption and physical activity that would not be possible if other sources of data are used. We tackle the measurement of IOp via two different methodological and philosophical approaches, one that is only concerned about inequalities between people that share the same circumstances (*ex-ante*), and another that focuses on inequality between people that exert equal effort (*ex-post*), the former concerned with the reward and the latter with the compensation principle. Furthermore, we measure these for several outcomes that account for different expressions of malnutrition such as obesity and underweight, but also undernourishment and anaemia. We find that *ex-ante* IOp has been persistent across the life course of all the individuals born between 1983 and 1988 and that inequalities in undernutrition have increased as individuals age, whereas inequalities in overnutrition have decreased as the cohort got older. Results also indicate that circumstances are the main driver of inequality in undernutrition across the lifespan. However, we do not find clear and consistent evidence that efforts account for most of inequalities in excess weight or adiposity measured through the BMI and WC. This evidence poses relevant questions regarding multiple aspects, for example, the idea of the dominant role of people's choices on obesity outcomes or the long-lasting effect of nutrition-related programmes for children that were and are currently implemented by the Federal government. Furthermore, this work contributes to the policy-making process in Mexico by identifying the accumulation of socioeconomic-related health inequalities during a life course of 30 years in malnutrition and revealing the transmission of inequities across generations. The investigation contributes to a much better and wider understanding of the evolution of inequalities. This piece of research innovates not only in looking at the dynamics of fair and unfair inequalities across the life cycle, but also in terms of malnutrition outcomes, and in proposing a technique that overcomes the lack of panel data that has hindered research about health inequalities across time. The remainder of this paper is divided into six sections. The following section presents two conceptual frameworks. First, the double burden of malnutrition is explained and second, IOp is conceptualised. The third section describes the empirical strategy and how this pseudo-birth cohort is constructed, as well as the approaches to the measurement of IOp. The fourth section explains the sources of data and describes the main variables of the analysis. The subsequent section shows the results of the analysis, and the final section closes by presenting conclusions and a discussion of the results.

2.2 Conceptual frameworks

2.2.1 Double burden of malnutrition

Even though malnutrition is the coexistence of under (*a lack of*) and over (*an excess of*) nutrition, many researchers and policy-makers have neglected this continuum and analysed these separately. Undernutrition is mostly related to the lack of micronutrients (vitamins and minerals) and overnutrition is conceived as an excess of macronutrients (proteins, carbohydrates, and fats), that leads to an excess of weight or adiposity. Furthermore, there seems to be a tacit idea that relates the two sides of malnutrition to specific age groups. For example, that undernutrition is mostly present in children and that obesity mostly happens among adults. This has been materialised in the nutrition-related policymaking of the past three decades in Mexico. Notwithstanding, empirical evidence has highlighted that a double burden can indeed manifest within populations, households, or individuals across the lifespan (WHO, 2017a). It could be the case that individuals that were exposed to different types of malnutrition during their childhood might be more likely to develop some sort of malnutrition later in life, but it could also be possible that at each point of life individuals present both under and overnutrition.

The double burden of malnutrition (DBM) is a worrying public health problem that many countries currently face. Its negative consequences are significant. First, it causes higher morbidity and mortality among populations. Second, undernutrition in the first stages of life can cause impairments to education, low capacity to resist diseases later in life and lower social and labour inclusion (Shrimpton et al., 2012). Third, it is costly for society. A study estimated that undernutrition costed, on average, 4.6% of the aggregated gross domestic product (GDP) of 11 Latin-American countries in 2017 (ECLAC, 2017). DBM represents a financial burden through higher associated health-care costs and lower labour productivity and, consequently, low economic growth and social development (WHO, 2017a).

The DBM is closely related to the familial context. There are at least three potential mechanisms of health transmissions across generations: 1) the latency model when some exposure over a specific period has a lifelong and irreversible effect on health that may be modified later; 2) the pathway model, when several biological and psycho-social intermediate factors between early life and adult health may all matter for health changes (Jacob et al., 2017) and, 3) the intergenerational transmission, where parental health is related to children's health (Trannoy et al., 2010). Thus, familial circumstances, behaviours and contexts (an obesogenic environment, for instance) are key factors for the future health status of individuals (Aitsi-Selmi, 2015; Crossman et al., 2006; Kral et al., 2010; McCormack et al., 2011; Reyes et al., 2004; Silveira et al., 2010). A better understanding of the origins and socioeconomic mechanisms behind the DBM is paramount to preventing and tackling the negative social and economic consequences of this phenomenon. The IOp framework offers a suitable approach to further this aim.

2.2.2 Inequality of opportunity

The (in)equality of opportunity framework was developed to distinguish between fair (legitimate) and unfair (illegitimate) sources of disparities. This implies that inequalities are not *per se* negative among societies. The vast literature on this topic has agreed upon two points. First, there are factors that individuals cannot control or choose (circumstances) and efforts that people exert based on their free will, in contrast to circumstances, efforts represent factors or choices that people can control, decide upon and therefore, be responsible for. Nevertheless, where to set the distinction between these two concepts has been a matter of debate, some argue that this differentiation can be made through the "responsibility-cut" (Jones, 2019; Roemer, 1998). Under this perspective, it is implicitly assumed that, to some extent, people are aware and conscious of the consequences of their acts. In the same vein, the distinction could be set according to a "legal age" (Arneson, 1989; Brunori, 2017; Jusot et al., 2019), that reflects the age at which individuals can consciously comprehend their acts and actions, and be accountable for their potential consequences. The differentiation between efforts and circumstances can be more clearly identified when agency and free will are attributed to individuals. This is where the responsibility cut according to a specific age is relevant. By setting this clear cut at 18 years of age, the empirical strategy for estimating ex-ante and ex-post inequality of opportunity assumes that people's efforts can only be observed after this age. The cut-off is not arbitrary and relies on the legal grounds embedded in several Mexican health laws, such as those related to individuals being socially allowed to buy alcohol and tobacco. These laws assume as well, that after that age, individuals can comprehend and face whatever consequences their choices may entail. By relying on this threshold, this investigation claims justification for public intervention since children and adolescents are vulnerable to their parents' or tutors' efforts and circumstances. Indeed, the measurement of IOp in children should include only circumstances since, by definition, children do not choose or decide upon their acts, and thus cannot be held responsible for their lifestyles or eating consumption decisions. At most, children's circumstances reflect their parents' efforts.

Second, there are two ethical principles reflected in the *ex-ante* and *ex-post* inequality measures: the reward and the compensation principles. The former demands that efforts exerted should be rewarded and respected when designing redistribution policies and the latter claims that inequality due to circumstances should be eliminated or compensated for (Jusot et al., 2019). The reward principle, associated with *ex-ante* IOp implies that an inequality measure should not reflect within-type inequality, while the compensation principle, related with the *ex-post* approach, implies that the inequality measure should fully reflect within-tranche inequality (Brunori et al., 2022). The IOp framework proposed by John Roemer offers a better comprehension of the interplay between circumstances and the role of mediating factors, such as efforts and choices, that people exert across different stages of the life course. The *ex ante* approach focuses on the measurement of people's opportunities *before* any effort is realised; thus, it concentrates on inequalities related to circumstances only (Davillas et al., 2020a). In contrast, the *ex post* considers heterogeneity in the level of outcomes within people that have exerted the same level of effort (Ramos et al., 2016). The application of this framework in tracking the DBM allows us to identify how inequalities potentially accumulate, transmit and reproduce across the life cycle and whether efforts play a mediating role in this process.

2.3 Empirical strategy

The analysis of the accumulation of inequalities across the life cycle would, ideally, require either household or individual panel data to track individuals over a long period of time. Unfortunately, such detailed longitudinal data for a significantly long-time horizon is not available either for Mexico or other low and middle-income countries. Instead, this study relies on repeated cross-sections of individuals and exploits matching and weighting techniques to construct a pseudo birth cohort to mimic life cycle data. The approach consists of observing a cohort of people born between 1983 and 1988 and following matched individuals across 30 years as the cohort has aged. Children that were newborns and up to five years old in 1988 are compared with matched older individuals into adulthood. In this way, conditional on matching, the study simulates the ageing of the initial cohort (see Table 2.1). For instance, individuals that were newborns in the 1988 survey would be 11 years old in 1999, 18 in 2006, 24 in 2012, 28 in 2016 and 30 years old in 2018. The use of matching and reweighting methods ensures that the six cross-sections can be regarded as representative samples of individuals from the same birth cohort at different points in time. This innovative way of dealing with the lack of longitudinal data, not only permits the study of the evolution of health inequalities over time but also guarantees that cross-sectional measures of inequality are comparable (in aggregate terms) over time since they represent the same underlying population represented by the birth cohort. Table 2.1 illustrates the study design. It shows how individuals included in all nutrition surveys in Mexico, and that were born between 1983 and 1988 are used to construct a pseudo-cohort to follow them from childhood into adulthood.

To ensure that the samples can be regarded as representative of individuals from the same birth cohort at different points in time, matching and weighting methods proposed by Blackwell et al. (2009) are followed. These techniques were originally designed for causal evaluation purposes, it is particularly useful when the identification strategy relies on observable characteristics and the treatment and control groups need to be balanced across covariates. We use the approach for group-balancing purposes. In what follows, the terms treatment and control groups refer to the different surveys. Formally, n represents a random sample taken from a population of N individuals, $n \leq N$. T_i is a variable that indicates whether individuals are present either in the treatment or control surveys. For our case, the treatment survey is the 1988 National Nutrition Survey (NNS) and all other surveys will be defined as controls, such that T_i=1 if i is in the 1988 NNS and T_i=0 if i belongs to the 1999, 2006, 2012, 2016 or 2018 surveys. Balance, defined as when covariate means across treatment and control cross-sections are statistically equivalent, is achieved by matching samples on observational data. Essentially selecting 1988 as a baseline (treated) group and separately matching observations from each of the other years to the baseline. The matching covariates, X, are time-invariant individual characteristics.

Once the matching variables have been defined, the following step is to create strata according to the matching covariates. Exact and many-to-one matching is used, such that multiple control individuals can be matched to a treated individual. Matching weights are calculated for each observation by dividing the number of treated by control observations in each stratum, adjusting by a normalisation factor (Porro et al., 2009), as:

$$W_{Cs} = \left(\frac{m_T^s}{m_C^s} \right) * \frac{m_{Cn}}{m_{Tn}} \tag{2.1}$$

Where m_T^s equals the number of treated observations T within strata s. Likewise, m_C^s is the number of control observations C within strata s. In the *normalisation factor*, $\frac{m_{Cn}}{m_{Tn}}$, m_{Tn} indicates the number of T in the n sample. The same applies to m_{Cn}. The individuals from the baseline will have weights equal to one and the control individuals in other years will be assigned a matching positive weight.

IOp is measured under both, the *ex-ante* and *ex-post* approaches assuming an ethical point of view where an age of responsibility cutoff differentiates between circumstances and efforts. For this analysis, this cut point is set at 18 years old. This age has been chosen based on the legal norms about the minimum age at which, in Mexico, it is permitted to buy and consume tobacco and alcoholic beverages, be able to vote and contract legal responsibilities such as: getting married, opening a bank account, eligibility for bank credit, etc. Furthermore, nutritional age restrictions are imposed due to paternalistic motivations to protect children who are deemed incapable of making rational decisions because they are unable to consider the future consequences of their actions.

Table 2.1: IOp across the lifespan. Empirical analysis design

	Survey year					
Year of birth	**1988**	**1999**	**2006**	**2012**	**2016**	**2018**
1988	>0 yo...	11 yo...	18 yo ...	24 yo...	28 yo...	30 yo
1987	1 yo...	12 yo...	19 yo...	25 yo...	29 yo...	31 yo...
1986	2 yo...	13 yo...	20 yo...	26 yo...	30 yo...	32 yo...
1985	3 yo...	14 yo...	21 yo...	27 yo...	31 yo...	33 yo...
1984	4 yo...	15 yo...	22 yo...	28 yo...	32 yo...	34 yo...
1983	$\leq$ 5 yo...	$\leq$ 16 yo...	$\leq$ 23 yo...	$\leq$ 29 yo...	$\leq$ 33 yo...	$\leq$ 35 yo...
	Children	**Adolescents**	**Adults**			
Survey respondents:						
	Preschool children: <5	Preschool children: <5	Preschool children: <5	Preschool children: <5	Preschool children:<5	Preschool children: <5
		School children:5-11	School children:5-11	School children:5-9	-School children:5-9	School children:5-9
		Adolescents: 12-19	Adolescents: 12-19	Adolescents: 10-19	Adolescents: 10-19	Adolescents: 10-19
	Women: 12-49	Women: 12-49	Adults >20	Adults >20	Adults >20	Adults >20
	Estimation of *ex-ante* inequality of opportunity					
			Estimation of *ex-post* inequality of opportunity			

Note: yo=years old. Columns two to seven show the expected age given the year of birth, as shown in column one. Survey respondents refer to the type of individuals that responded to the survey in each year. The vertical red line sets the responsibility-legal age. The last two rows indicate the estimation approach to IOp given this age cut-off.

2.3.1 Measuring inequality of opportunity

We measure *ex-ante* IOp at different points of age, following the direct, mean-based and parametric approach proposed by Ferreira et al. (2011) and that has already been applied to health by Davillas et al. (2020a). We also measure *ex-post* IOp following the approach proposed originally by Jusot et al. (2013) and recently applied to biomarkers by Carrieri et al. (2020). This latter approach is only estimated for individuals above the "responsibility-cut" age.

Roemer's benchmark model applied to health assumes that a health outcome (y_i) of an individual i is a function of their circumstances, C, their efforts, E, and other random factors, such as "luck", u_i. In this model, individuals that observe the same circumstances belong to the same *type*, whereas those that exert the same effort belong to the same *tranche*. Individuals that share the same circumstances and efforts belong to the same *cell*. In a non-parametric approach, Roemer proposes to split the distribution of effort into quantiles to make the degree of effort exerted by individuals of different types comparable. Individuals that exerted the same degree of effort and therefore belong to the same quantile (q) within each type belong to the same cell as well. An important aspect of the model is that the distribution of effort within each type is a circumstance by itself since it is beyond the individual's control (Rosa Dias, 2009). The model allows efforts to be dependent on individual circumstances together with factors that are beyond people's circumstances, v_i.

$$y_i = h(C_i, E(C_i, v_i), u_i) \tag{2.2}$$

Assuming additive separability and linearity in $h(.)$ and $E(.)$ a system of equations can be written in the following structural form:

$$y_i = \alpha_0 + \alpha_1 C_i + \alpha_2 E_i + u_i \tag{2.3}$$

$$E_i = \delta_0 + \delta_1 C_i + v_i \tag{2.4}$$

In Equation (2.3), α_1 and α_2 are coefficients that respectively capture the direct effect of circumstances and efforts on the outcome y for individual $i = 1, ..., N$. In Equation (2.4), δ_1 represents the indirect effect of circumstances on efforts. It is worth noticing that Equation (2.4) can be rearranged to show efforts purged from the effect of circumstances such that,

$$v_i = E_i - \delta_0 - \delta_1 C_i \tag{2.5}$$

which is equivalent to

$$v_i = E_i - \mathbb{E}(E_i|C_i) \tag{2.6}$$

Thus, the estimator of v_i is

$$\hat{v_i} = E_i - \hat{E_i} \tag{2.7}$$

The model also assumes that $h(.)$ is continuous and strictly increasing in C and E.

Ex-ante **IOp**

This type of IOp focuses on measuring the role of circumstances on people's outcomes before efforts are exerted. Thus, in the *ex-ante* approach effort is unobservable. In practice, we have the mean-based, parametric and reduced form to measure *ex-ante* IOp by inserting Equation (2.4) into (2.3) and arranging the terms, such that:

$$y_i = \beta_0 + \beta_1 C_i + \epsilon_i \tag{2.8}$$

Where y_i are the health outcomes for individual i. $\beta_0 = (\alpha_0 + \alpha_2\delta_0)$ is the intercept, $\beta_1 = (\alpha_1 + \alpha_2\delta_1)$ captures the total contribution of circumstances, reflecting the direct effects of circumstances on the outcomes and the indirect effect of circumstances through efforts. $\epsilon_i = (\alpha_2 v_i + u_i)$ depicts the error term that captures random factors not captured by circumstances or efforts such as luck, lack of talent, motivation, physical impediments, etc. Hence, from Equation (2.8) the total effect of circumstances, which comprises the direct and indirect effect of circumstances through effort, is estimated.

The mean-based approach assumes inequality neutrality. Given that effort is not observed and that the stock of health monotonically increases with effort, the *ex-ante* approach also assumes that once types are fixed, effort is the only determinant of health. Thus, within each type, those individuals at the q^{th} quantile of the outcome distribution, on average, also belong to the q^{th} quantile of the effort distribution (Rosa Dias, 2009).

Measuring *ex-ante* IOp is a two-step procedure. The first is to estimate Equation (2.8) to obtain a counterfactual distribution of the outcome *if* no differences in outcomes arise as a consequence of having different circumstances (Davillas et al., 2020a). In practical terms, this corresponds to the predicted values of health. The second step is to plug the predicted values into an inequality measure (Ferreira et al., 2011). The way to estimate Equation (2.8) and the inequality measure to use, depends not only on the desired properties of the inequality measure but also on the type of health outcome variable. Chávez -Juárez et al. (2014) argue that when the outcome variable is continuous on an inherent scale, the best choice is to use the mean-logarithmic deviation (MLD). For binary variables that are scale-invariant, the dissimilarity index (D-index) is preferred. We use the D-index as an inequality measure when estimating IOp at specific clinical thresholds for under and overnutrition using the dichotomised version of our outcomes and the MLD when estimating IOp across the continuous distribution of our outcomes. Thus, we use logit and linear models to estimate Equation (2.8) for binary and continuous outcomes, respectively. The MLD measures the deviation of the expected level of health outcome from the group's expected average. Smaller values reflect lower levels of IOp. MLD is defined as:

$$MLD(\hat{y}) = \frac{1}{N}\sum_{i=1}^{N} ln\frac{\bar{\hat{y}}}{\hat{y}_i} \tag{2.9}$$

where $\hat{y_i} = \mathbb{E}(y|C_i)$. Absolute inequality is obtained when the counterfactual distribution of health outcomes conditioned on circumstances is plugged into the MLD, such that:

$$\theta_a = I_0(\hat{y_i}) \tag{2.10}$$

and relative IOp is the ratio of the absolute level of inequality concerning the overall inequality, as:

$$\theta_r = \frac{I_0(\hat{y_i})}{I_0(y_i)} \tag{2.11}$$

Relative inequality is zero when equality is observed, and positive values depict an unequal distribution of the outcomes.

The D-index is an absolute measure that focuses on the dissimilarity of the level of health for groups defined by their circumstances compared with the average level of health of the population. Another way to interpret the index is as a weighted mean of the absolute differences of the estimated outcome, from the overall outcome average. If equality exists, D=0 (Paes de Barros et al., 2008). The D-index is defined as:

$$\theta_a = D(\hat{y}) = \frac{1}{2N\bar{\hat{y}}}\sum_{i=1}^{N} |\hat{y}_i - \bar{\hat{y}}| \tag{2.12}$$

where $\hat{y_i} = \mathbb{E}(y|C_i)$. In this case, y is a binary variable. Thus,

$$Prob\{y_i = 1\} = (e^{\beta_0+\beta+\epsilon_i}) \cdot (1 + e^{\beta_0+\beta_1 C_i+\epsilon_i})^{-1} \tag{2.13}$$

and

$$Prob\{y_i = 0\} = (1 + e^{\beta_0+\beta_1 C_i+\epsilon_i})^{-1} \tag{2.14}$$

with $\bar{\hat{y}} = \frac{1}{N}\sum_{i=1}^{N} \hat{y_i}$.

***Ex-post* IOp**

For individuals above the "age of responsibility" threshold, it is assumed that health outcomes are explained by circumstances, efforts, and other random factors. In this approach, efforts are assumed to be mediating factors that lie in the pathway between circumstances and outcomes. Since efforts are observable, it is possible to disentangle two different effects: i) the total effect of circumstances, which can be further decomposed into a direct and indirect effect and, ii) the indirect effect of efforts on outcomes.

Since the efforts exerted depend on people's circumstances, the former and the latter are correlated. To deal with this issue, *ex-post* IOp is calculated via a two-stage model. First, Equation (2.4) is estimated to remove the influence of circumstances on efforts, we estimate this equation using OLS models. By doing this, we can obtain Equation (2.7), which represents the *purged* level of effort. In the second stage, the outcome variable is regressed against the vector of circumstances and the isolated level of effort (Jusot et al., 2013).

Since $E_i = \hat{v}_i + \hat{E}_i$, we can re-write $h(.)$ as:

$$y_i = \alpha_0 + \alpha_1 C_i + \alpha_2(\hat{v}_i + \hat{E}_i) + u_i \tag{2.15}$$

and by separating terms, we obtain the total effect of circumstances, which can be further decomposed into a direct and an indirect effect, and the direct effect of efforts.

$$y_i = \alpha_0 + \underbrace{\underbrace{\alpha_1 C_i}_{\text{direct effect of } C} + \overbrace{\alpha_2 \hat{E}_i}^{\text{indirect effect of } C}}_{\text{total effect of } C} + \underbrace{\alpha_2 \hat{v}_i}_{\text{direct effect of } E} + u_i \tag{2.16}$$

A simplified form of Equation (2.16) can be written as:

$$y_i = \gamma_0 + \gamma_1 C_i + \gamma_2 \hat{v}_i + u_i \tag{2.17}$$

Where $\gamma_0 = (\alpha_0 + \alpha_2 \hat{\delta_0})$, $\gamma_1 = (\alpha_1 + \alpha_2 \hat{\delta_1})$ and $\gamma_2 = \alpha_2$. In Equation (2.17), y_i represents the health outcomes, undernutrition as well as overnutrition outcomes defined at clinical cut-off points, as well as a binary outcome that explicitly measures malnutrition by combining under and overnutrition outcomes defined by clinical thresholds, for individual i that is above the legal cut-off of 18 years old. γ_0 is the constant term; C is the vector of circumstances and γ_1 captures the total contribution of circumstances on outcome y and γ_2 is a vector that captures the direct contribution of efforts. The relation between the *ex-ante* and *ex-post* approaches is the equivalence of the estimators β_1 in Equation (2.8) and γ_1 in Equation (2.17), both represent the total effect of circumstances on the outcome y. The former in the *ex-ante*, the latter under the *ex-post* approach.

For the *ex-post* case, the variance is used as an inequality measure. To ease the interpretation of the results, the level of inequality is disentangled between the total effect of circumstances and the direct effect of efforts. For this, we rely on previous work developed by Deutsch et al. (2018) who used a Shapley-inspired approach to decompose the variance, in our case the McFadden's R-squared[1]. y could be binary and take the value of 1 if the individual has some sort of malnutrition and 0 otherwise. When estimating Equation (2.17) via a logit model,

$$Prob\{y_i = 1\} = (e^{\gamma_0 + \gamma_1 C_i + \gamma_1 \hat{v_i} + u_i}) \cdot (1 + e^{\gamma_0 + \gamma_1 C_i + \gamma_1 \hat{v_i} + u_i})^{-1} \tag{2.18}$$

and

$$Prob\{y_i = 0\} = (1 + e^{\gamma_0 + \gamma_1 C_i + \gamma_1 \hat{v_i} + u_i})^{-1} \tag{2.19}$$

Thus, the Shapley-inspired decomposition is based on the idea of comparing the indicator of goodness-of-fit, McR^2, when including all circumstances and efforts versus another model in which only efforts are included, for example. The likelihood ratio that corresponds to the logit model is:

$$LL_M = LL(C_i \neq 0, \hat{v}_i \neq 0) \tag{2.20}$$

where $\neq 0$ means that all coefficients in the model are unrestricted. C_i denotes the vector of circumstance for individual i. $\hat{v}_i$ represents the vector of efforts for individuals i. If the vector of

[1] $McR^2 = 1 - \frac{LL_M}{LL_0}$ where LL_M corresponds to the model value of the log-likelihood and LL_0 the log-likelihood when only the constant term is introduced (Deutsch et al., 2018).

efforts $\hat{v}_i$ is not included in the model:

$$LL_C = LL(C_i \neq 0, \hat{v}_i = 0) \tag{2.21}$$

And similarly, if the vector of circumstances C_i is not included, the LL can be written as:

$$LL_v = LL(C_i = 0, \hat{v}_i \neq 0) \tag{2.22}$$

Using the Shapley decomposition approach, the marginal contribution of circumstances to the actual likelihood ratio is calculated as:

$$TCC = 0.5(LL_C) + 0.5(LL_M - LL_v) \tag{2.23}$$

and the contribution of efforts as:

$$DCE = 0.5(LL_M - LL_C) + 0.5(LL_v) \tag{2.24}$$

We can then check that, $TCC + DCE = LL_M$. Even though Deutsch et al. (2018) proposed this decomposition for the McFadden R-squared for logit models, the approach can also be used to decompose the R squared in OLS models. Important is to note that such decomposition should not be understood as causality, but only to show the relative importance of circumstances and efforts. This is because unobservable determinants of nutrition-related outcomes are likely to be correlated with the observable circumstances (Ferreira et al., 2011). Furthermore, while the *ex-ante* approach takes all efforts into account (both observed and unobserved), there might still be unobserved circumstances. β_1, in the reduced form, estimates both the direct effect of circumstances and their indirect effect through efforts. Since there might be omitted circumstances, we get a lower bound of ex-ante IOp. Due to the same reason, and despite circumstances being pre-determined and entirely exogenous to people, these estimates cannot be seen as causal. Besides this, the *ex-post* approach also suffers potential omitted variable bias due to unobserved efforts.

2.3.2 Additional analyses

As illustrated in Figure 2.1, both *ex-ante* and *ex-post* IOp is calculated for each survey year. This is achieved by estimating models depicted in Equations (2.8) and (2.17) and using the weights calculated in the matching and re-weighting exercise. The main outcomes are binary indicators of specific health outcomes, for example, under and overnutrition and a measure of malnutrition, defined according to clinical thresholds. The main approach is mean-based and assumes inequality neutrality. However, this approach is limited given that malnutrition can be found at the bottom, as well as at the upper parts of some of the outcome distributions, such as BMI and WC.

By relaxing inequality neutrality and allowing for inequality aversion, we explore the role of circumstances alone and circumstances and efforts in *ex-ante* and *ex-post* IOp, respectively. Thus, we measure IOp at the 10^{th}, 25^{th}, 50^{th}, 75^{th}, 90^{th}, 95^{th}, 99^{th} percentiles of the BMI and WC distributions. Table B.1.2, in the Appendix, summarises the outcomes used in each approach. To ease the flow of the paper, in the main text we present the mean-based estimations of IOp using the D-index as the inequality measure for the *ex-ante* approach and the variance for the *ex-post*.

Given the structure of the data, the set of circumstances included in each model differs according to the year of analysis. Thus, we present the results of IOp using a vector of common circumstances across all survey years (sex, ethnicity, running water inside the house, household living standards and geographic region). *Ex-ante* IOp was not calculated in 2016 due to the restrictions imposed by the survey. Equally, *ex-post* IOp was not estimated in the 2012 and 2016 survey years due to sample sizes being too small to run the models and the lack of effort variables. Specifically, the number of potential observations in the 2012 survey were: n=34 for the DBM outcome, n=35 for the HB, n=89 for the BMI, and n=85 for the WC models, respectively. Considering the number of circumstances and effort variables, there was likely to be insufficient degrees of freedom for estimation. Furthermore, we did not estimate IOp using the 2016 survey because the 2016-survey was a mid-way survey and rich data on food consumption and physical activity were not collected. Thus, there were no effort variables available.

Given biological (Thurstans et al., 2020), social (Matthews et al., 1999) and cultural reasons (Kanter et al., 2012), it is expected that the contribution of different factors to nutrition-related outcomes differ across sex. Thus, we further explore *ex-ante* and *ex-post* IOp in all outcomes separately for men and women. Additionally, we also distinguish the contribution of sex, as an independent source of inequalities, in the *ex-post* decomposition. Results from these analyses are shown in the Appendix, Figures B.9.1 to B.9.8.

2.4 Data

2.4.1 Sources of data

Data from the 1988 and 1999 NNS and the National Health and Nutrition Surveys (ENSANUTs) from 2006, 2012, 2016 and 2018 are used. These are cross-sectional and represent all the health and nutrition surveys carried out in Mexico. The 1988 survey was the first-ever national nutrition survey conducted in Mexico. This survey collected data from more than 13,000 households. The study population were children under five years and women between 12 and 49 years of age. The resulting sample was representative at the national level (Consejo Nacional de Evaluación de la Política Social, 2010; Resano-Pérez et al., 2003). The second survey, the 1999 national nutrition survey, collected data between October 1998 and March 1999. The sample consists of almost 18,000 households, that nationally represented areas with less than 2,500 inhabitants, with 2,500 to 14,999 inhabitants and areas with more than 15,000 inhabitants. The study populations were children under 5 years of age, school-age individuals between 5 and 19 years of age and women[2] aged 12 to 49 (Resano-Pérez et al., 2003).

Ensanut 2006 collected data from around 44,500 households. It is representative at the national level, as well as urban (> 2,500 people) and rural areas (<2,500 people). The survey respondents were children under five years of age; children of school age 5-11; adolescents (12 to 19 years of age) and adults, men, and women over 20 years of age (Gustavo et al., 2006). Ensanut 2012 is as well a national representative survey. The sampling design was probabilistic, multi-stage and stratified. Data collected held information from 50,528 Mexican households. The survey respondents were children less than five years old, school children (5 and 9 years old), adolescents (age between 10 and 19 years old) and adults older than 20 years old (Romero-Martínez et al., 2013). Ensanut 2016 is a mid-way survey[3] with smaller sample size, compared with the Ensanut 2012 (9,479 households (Romero-Martínez et al., 2017)). Some of the questions asked in the 2012 survey were not included in the 2016 survey. Ensanut 2016 had a different sampling design. Data from Colima and Oaxaca States were not collected (Romero-Martínez et al., 2017). Additionally, data on food consumption were not as detailed as in the 2012 wave. Finally, Ensanut 2018 is a national representative survey. The sampling design was probabilistic, multi-stage and stratified by rural and urban areas. Data from 50,654 households were collected and survey respondents as well children less than five years

[2] It is worth noticing that only adult women were targeted in the 1988 and 1999 national nutrition surveys.

[3] Since 2006, Ensanut was thought to be administered every six years. Notwithstanding, as a response to the accelerated increase in the prevalence of overweight and obesity, it was decided to conduct a *mid-term* survey between 2012 and 2018 to assess the nutritional status of the population (National Institute of Public Health, 2016)

old, school children from 5 and 9 years old, adolescents aged 10 to 19 years and adults older than 20 years old (Romero-Martínez et al., 2019). The sample units across all the surveys were households. A household is defined as a group of people, related by kinship or not, who usually sleep in a house under the same roof, benefiting from a common income, from either one or more of the household members. The key respondents were those individuals that resided in households at the time of the study. This impacts the analysis in two ways: 1) in some cases, it is not possible to identify the familial link and given that, 2) the set of circumstances included in each model is different in each cross-section according to the availability of data.

2.4.2 Key variables

2.4.3 Matching variables

The matching covariates, X, are time-invariant individual characteristics, such as year of birth, sex and geographical region where individuals lived. The choice of matching variables is data-availability driven. The selection of region, instead of geographical State where people live, as a time-invariant characteristic, is based on the evidence that the internal migration in Mexico occurs mainly between regions (Rangel Garrocho et al., 2014; Sobrino, 2010). This assumes that even though individuals could migrate, this migration mainly occurs within the geographical region and not across them. Even though this assumption is not testable in all surveys, in Ensanut 2018 people were asked about the State where they were born and the State where they lived when the survey took place. Results show that 82% of the respondents said they were born in the same State where they were living at the moment of the survey. 17% said they were born in another State, but 97% were born in another State of the same geographical region. Thus, these results give confidence for our assumption to hold. A description of the distribution of matched individuals by survey years and birth cohort is shown in Table B.3.1, in the Appendix.

Outcomes

Malnutrition is a spectrum that can comprise under and overnutrition. The concept of DBM, which appeared at the beginning of the 1990s, refers to the coexistence of over and undernutrition that can occur within populations, households, and individuals. Thus, it is possible to have a DBM at the individual level, for instance, a person overweight and that also presents a lack of nutrients. Throughout the life course, the way to measure nutrition-related outcomes cannot be consistently the same. It is well known that weight and height vary with age, and this is particularly important during childhood. Classically, HAZ, WHZ and WAZ are used in the public health literature to assess the nutritional status of children. The Z-scores benchmark height and weight differences for all children of the same age to allow comparability. Thus, given the lifespan approach adopted, it is not possible to use the same outcomes across all age groups.

We use low HAZ as a proxy for stunting, low WHZ for wasting, low WAZ to capture undernutrition and BMI-for-age as a proxy of overnutrition when analysing child data. This refers to the 1988 survey, where individuals were children under five years of age. For the rest of the cross-sections, the level of haemoglobin (HB) in the blood is used to account for undernutrition since low levels of haemoglobin in the blood are closely related to anaemia, defined as a deficiency of iron, folate and vitamin B (WHO, 2012). The body mass index and waist circumference (WC) are used as outcome variables that capture overnutrition in the form of excess weight or excess body fat and excess or central adiposity in adolescents and adults.

Undernutrition is operationalised in the following ways. For the child-level analysis (1988 cross-section), HAZ, WHZ and WAZ are used as outcome variables. We discretised the variables such that 1 depicts if the outcome is lower than -2 z-scores and 0 otherwise [4]. For the adolescent and adult-level analysis (1999 to 2018 cross-sections), the level of haemoglobin is used as a measure of undernutrition. This outcome is used both in a binary and continuous way. This aims to measure IOp using the clinical cut-off points, but also to allow IOp measurement across the whole distribution. For the binary case, the variable takes the value of 1 if hb<12g/dl[5] for females and <13 g/dl for males, and 0 otherwise (WHO, 2012). Overnutrition is defined when a child has a BMI z-score above 2 and when an adult observes a BMI above 25 kg/m^2 or WC above 80cm and 90cm (for women and men, respectively). Furthermore, an outcome that explicitly accounts for malnutrition (under and overnutrition) is constructed and defined as a binary variable that takes the value of 1 when a child has a z-score for HAZ or WHZ or WAZ below -2 *and* BMI above 2. Likewise, for adolescents and adults, malnutrition is defined when they present a BMI above 25 kg/m^2 or WC>80 or 90cm according to sex *and* HB<13 g/dl.

Data on height, weight and blood samples were taken and measured by specialised and trained staff by the National Institute of Public Health (INSP), in Mexico. Haemoglobin (g/dl) was adjusted for altitude and smoking behaviours[6]. For the analysis, implausible biological values for

[4] The WHO defines moderate to severe undernutrition when height-for-age is <-2 z scores (WHO, 2020)

[5] g/dl means grams per decilitre

[6] Raw values of haemoglobin were adjusted in the following way. For altitude: no change if altitude<1,000m; -0.2 if $\geq 1,000$m but <1,500m; -0.5 if $\geq 1,500$m but <2,000m; -0.8 if $\geq 2,000$m but <2,500m; -1.3 if $\geq 2,500$m but <3,000m and -1.9 if altitude $\geq 3,000$m but <3,500m. For smoking behaviours: -0.03 if the the individual smokes

z-scores are excluded, as follows: <-5 and >+3; for WHZ <-4 and >+5 and <-5 and >+5 for WAZ and BMIZ (O'Donnell et al., 2007). For BMI values below 10 kg/m^2 and above 59 kg/m^2 are excluded. For WC, measurements below 51cm and above 190 cm are excluded. Information from pregnant women is not used in the analysis. We also exclude implausible values for adjusted haemoglobin such as those below 4g/dl and above 20 g/dl (Sullivan et al., 2008).

Circumstances

Framed within the social, political, and economic context previously described, we define circumstances as those factors beyond an individual's control and that are potential sources of illegitimate inequalities. These factors are categorised as those related to individual factors; family-related characteristics; household-level factors and geographical characteristics. The set of circumstances included has been chosen based on normative criteria according to those factors that have been socially defined as illegitimate sources of disparities. We next set out the rationale for the inclusion of specific circumstances as variables. A succinct list of outcomes, circumstances and efforts used in the analysis is found in Table B.1.1 in the Appendix.

Ethnicity, which refers to the ethnic background of the person. We used the official definition of ethnicity according to whether a person declares to speak an indigenous language. **Mother's health insurance**, this circumstance was chosen for two reasons. First, it is a proxy for preventive healthcare utilisation. Second, even though access to health is a fundamental right stated in the Mexican Constitution, specifically in Article fourth, accessibility to the health system in Mexico is heavily conditioned on accessing the labour market. **Mother's BMI (kg/m^2)**, this circumstance is a proxy indicator of nutrition in the household (taking into account that women have traditionally taken care of the preparation of food in the household). This circumstance is framed in the literature related to the relationship between parental health and children's health outcomes. Given that in 1988 and 1999 the surveys focused on women, only the BMI of mothers is obtained. Another circumstance included is **mother's anaemia**, this circumstance is also a proxy indicator of the nutritional status of the mother along with **parental diabetes**, this variable accounts for the health situation of the parents and aims to reflect some indirect transmission of health conditions and eating behaviours since 90% of the cases of diabetes type II in Mexico are closely related to overweight and obesity (Dávila-Torres et al., 2015; Health, 2010). We also include **parent's education**, when the information was available.

up to one pack of cigarettes per day (20 cigars per package); -0.05 if $\geq$ 1 but <2 packs and -0.07 if $\geq$ 2 packs of cigarettes (WHO, 2012).

To account for household conditions, we include **running water available in the household** as a circumstance. Even though this public service was established as a constitutional right, as part of the right to a healthy environment (Article fourth, 1999), there are still evident gaps. According to the 2015 National Household Census, 94.4% of the Mexican population had running water in their houses. Nevertheless, there are marked differences across the States. In the Southern States, the percentage of households with availability of running water is, on average, 85% (INEGI, 2015). Hence, to have running water in a household cannot be taken for granted and represents a potential source of illegitimate health disparities. We also include **household living standards**, this circumstance is estimated using principal components techniques. The household asset index considered information reported by the head of the household about the physical characteristics of the house, for instance, material from which the floors, walls and roofs are made; number of rooms, whether the house has latrines or toilets, etc. The index also included data about the ownership of durable assets in the house, such as radio, television, fridge, telephone, car, computers, washing machine, microwave, air conditioner, etc. Using the polychoric principal-component analysis method and following Basto-Abreu et al. (2018), asset indices are estimated. A single component, which in all cross-sections explained more than 50% of the variation, was extracted. The index was then categorised into quartiles.

To account for potential geographic factors driving inequalities, we include the **State deprivation level**, this is a weighted index that measures social deprivation at the State level and takes into account characteristics such as the percentage of the population older than 15 years old and deemed illiterate, the population aged 6 to 14 who do not attend school, households with individuals aged 15 to 29 that have less than 9 years of education; population older than 15 years with incomplete basic education; population without health insurance; households with no floor; average occupants per bedroom; households without a toilet; piped water from the public network; sewage; electricity; washing machine or fridge. This index is estimated by the National Council for the Evaluation of Social Development Policy (CONEVAL) and the data come from the National Count of Population and Housing censuses (CONAPO), we used 1990, 1995, 2005, 2010 and 2015 indices. The State deprivation index aims to capture the *geography of opportunity*, a concept that describes how the area and geographical space where people live condition access to opportunities (Rosenbaum, 1995).

We also include the variables used in the matching exercise **sex** and **geographic region** where individuals lived when the survey was undertaken, since these are also circumstances[7].

Efforts

Efforts are those factors that lie within the sphere of individual responsibility and are subject to choice, such as lifestyles (Jusot et al., 2019), eating patterns and human capital investments that individuals above the age of responsibility decide to adopt/acquire. Rich data about dietary intake was collected using a food frequency questionnaire in Ensanut 2006, 2012 and 2018. This questionnaire included 101 different foods and beverages. For each food item, data about intake according to the number of days in the week, daily frequency of consumption, portion size and the number of portions consumed were collected[8]. With this rich data, we use factor analysis to identify dietary patterns for everyone (Denova-Gutiérrez et al., 2016). Three factors that accounted for approximately 35-50% of the total variance were extracted. We check also for the adequacy of using factor analysis using the Kaiser-Meyer-Olkin(kmo) test. The KMO coefficient was, on average, 0.85 which indicates that the sample is adequate for using the method. We characterise the three factors by looking at the loading of a given food to each factor and their correlations. The first pattern was characterised by grouping low-nutritious and high-energy food [9], the second pattern included high-nutritious food[10], and the third factor food items such as legumes and maize-based products[11]. To aid the interpretation of the dietary patterns, we re-scale them to range between 0 and 100.

For our analysis, efforts comprise: **dietary patterns** that includes three food groups, **nutritional supplements consumption**, this is a binary variable that takes the value of 1 if individuals reported that they consumed nutritional supplements, as part of their diet. We also include **daily hours of vigorous and moderate physical activity**, this variable is the sum of the self-reported daily hours doing vigorous physical activity, defined as those activities that take more than 6 metabolic equivalents (METs), such as aerobics, cycling fast, lifting heavy things, digging, doing farm work, etc. And moderate physical activity is defined as activities that use 3 to 6 METs, for example: carrying light things from one place to another, cleaning heavily, cycling at a regular pace, recreationally playing sports, etc. The number of hours was bounded from 0 to

[7] The 32 Federal States of Mexico were grouped in six regions: Northwest: Baja California, Baja California Sur, Sinaloa and Sonora. Northeast: Coahuila, Nuevo León, Tamaulipas, Chihuahua, Durango, Zacatecas and San Luis Potosí. West: Aguascalientes, Colima, Guanajuato, Jalisco, Michoacán, Nayarit and Queretaro. Centre: Mexico City, State of Mexico, Hidalgo, Morelos, Puebla and Tlaxcala. South: Guerrero, Oaxaca, Chiapas, and Veracruz. Southeast: Campeche, Quintana Roo, Tabasco and Yucatán.

[8] We classified the 101 food items into five different groups. See the Appendix Tables (B.5.1) and (B.5.2) for further information about the items and food groups.

[9] Whole-fat dairy, fast food, sweetened beverages, sweets and red meat

[10] Fruits, vegetables, poultry, fish, cereals

[11] Beans and maize-based products are the staples of the traditional and pre-Hispanic diets

16 hours in a day.

As risky health behaviours, we include **alcohol consumption**, this binary variable depicts if individuals consume alcohol above the number of daily units recommended, that is more than three units for women and more than four units for men[12]. Another risky behaviour included is **smoking**, this binary variable depicts whether an individual smokes tobacco frequently, meaning daily or regularly.

2.5 Results

2.5.1 Description of the sample and balancing weights

Figure 2.1 shows the distribution of the matching weights across the different surveys by year of birth. The empty space for those born in 1983 and surveyed in 1988 reflects the presumption that the 1988 survey collected information after December 1988, thus there were no children born in 1983 in this survey. As the treatment group are those individuals in the 1988 survey, their weights are equal to one. The highest range is found in the 2006 cross-section. Note that men living in the Northwest region of Mexico were not surveyed in the 2006 dataset, and hence were excluded from the analysis. See Table B.2.1, in the Appendix, for more details about the results of the matching exercise.

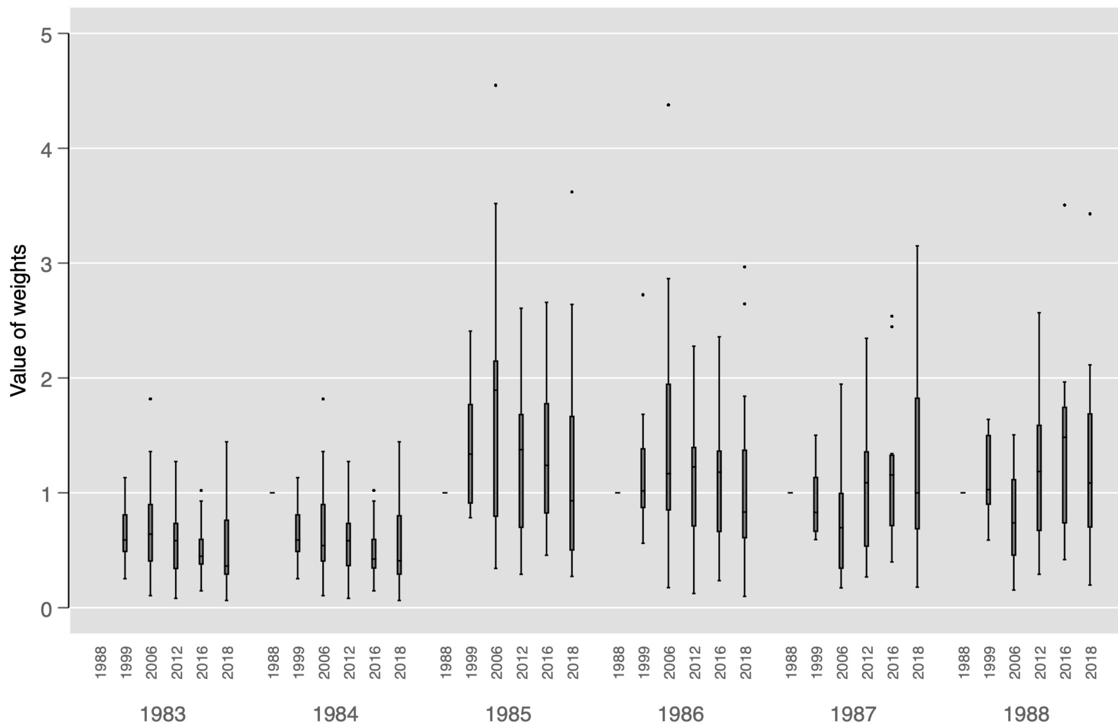
2.5.2 Description of key variables

Outcomes

Table 2.2 shows the description of the dichotomised outcomes according to the clinical cut-off points for malnutrition in terms of stunning, wasting and underweight, as well as overweight and obesity. Specifically, in 1988 23% of children below five years old were stunted, 6% wasted and 10% underweight. For overnutrition, 8% were overweight and 3% suffered a double burden of malnutrition. The table also shows the evolution of the cohort across time for undernutrition proxied by the presence of anaemia, overnutrition proxied by excess weight and adiposity, as well as malnutrition. In terms of undernutrition, the proportion of individuals with anaemia decreased from adolescence (1999 survey) to young adulthood (2012 survey). But, as the cohort get older, the proportion of anaemia among individuals in the 2016 and 2018 surveys doubled.

[12]Examples of one unit: one standard glass of 13%-level-of-alcohol wine; 25 millilitres (ml) of 40°-spirit; 250 ml of 4%-level-of-alcohol beer. One unit is also equivalent to 10ml or 8 grams of pure alcohol

Figure 2.1: Distribution of matching weights by cohort's year of birth



Contrary to the non-linear trend in the proportion of individuals with anaemia, the proportions of individuals with excess weight and adiposity have been continuously increasing over time. Table 2.2 shows that in 1999, the proportion of adolescents categorised as having an excessive weight was 11%, while in 2018, it was 75%. When looking at the proxy for central adiposity (WC), the proportion increased from 10% during the cohort's adolescence, to 77% during adulthood. The evolution of malnutrition is striking, the proportion of individuals with an excess of body fat or central adiposity and that also had anaemia increased from 3%, when individuals were aged 11 to 16 years old, to 41% when they were in their 30 to 35 years of age.

While Table 2.2 describes the outcome variables given the clinical cut-off points, Figure B.4.1, in the Appendix, shows a more detailed description of the distributions of the outcomes. Overall, both graphs show a similar story. Over the whole period, undernutrition (HB) increased from an average of 13.2 to 13.8 g/dl, with the trend being non-linear. For overnutrition (BMI and WC), Figure B.4.1 shows an increasing trend. On average, BMI and WC rose as people aged. Furthermore, BMI increased considerably from 20 kg/m2 (clinically considered as *normal*) to 28.2 kg/m2, which lies in the cut-off of *overweight*. The same applies to the WC, which rose from 67 cm in 1999 to 93.4 cm in 2018.

Table 2.2: Description of the binary outcomes across the lifespan

	Expected age of the Cohort					
	0-5	11-16	18-23	24-29	28-33	30-35
Survey year	1988	1999	2006	2012	2016	2018
Stunting						
Proportion	0.23					
N	6,003					
Wasting						
Proportion	0.06					
N	6,077					
Underweight						
Proportion	0.10					
N	6,228					
Anaemia (HB)						
Proportion		0.23	0.11	0.09	0.17	0.21
N		1,375	5,386	1,121	488	2,078
Overweight children						
Proportion	0.08					
N	6,101					
Excess weight (BMI)						
Proportion		0.11	0.39	0.58	0.70	0.75
N		3,256	5,636	1,198	1,059	2,112
Excess adiposity (WC)						
Proportion		0.10	0.51	0.59	0.73	0.77
N		2,101	2,701	1,118	984	2,019
Malnutrition children						
Proportion	0.03					
N	5,701					
Malnutrition*						
Proportion		0.03	0.08	0.10	0.29	0.41
N		978	3,354	530	175	827

Notes: Matching weights used. Stunting=Height-for-age (HAZ) below -2 Z scores. Wasting=Weight-for-height (WHZ) below -2 Z scores. Underweight=Weight-for-age (WAZ) below -2 Z scores. Anaemia: HB= 1 if Haemoglobin<13(g/dl) for men and <12 for women. Overweight children=Body mass index (BMI) Z score above 2. Excess weight: 1 if BMI>25kg/m2. Excess adiposity= 1 if WC>80cm for women and >90 for men. Malnutrition children=those that observe HAZ or WHZ or WAZ below -2 Z scores and BMI above 2 Z scores. Malnutrition*=those individuals that observe BMI above 25kg/m2 or WC>80cm (women) or WC>90cm (men) and HB<12g/dl(women) or <13 (men)

Circumstances and efforts

Table 2.3 shows the description of the sample in terms of individuals' circumstances and efforts. In all survey years, 51% of the individuals are males, and in general, the proportion of non-indigenous people is between 94% and 89%. In 1988, the BMI of children's mothers was on average 23.84 kg/mts^2 and 12% of the mother's cohort were anaemic. This table also shows that around half of the sample's mothers had health insurance when individuals were children. Regarding the mother's education, in 1988, 12% did not have any level of education, 59% had achieved primary school as the highest level of education, and approximately 10% had education above high school. In 1999, the proportion of mothers with no education, as well as the proportion of mothers with education above high school slightly increased. Most mothers had a level of education up to primary school. Table 2.3 also shows the proportion of parents that had not been clinically diagnosed with type II diabetes. Here, parental diabetes means that at least one of the parents did not have the condition. In 1988, the proportion was high (98%). Nevertheless, this proportion has been diminishing, reflecting the rapid increase in the prevalence of type II diabetes in the Mexican population.

Table 2.3 also depicts the household and municipal characteristics. The proportion of households with running water varies from a low of 21% in 2006 to a high of 62% in 2012. In 1988 many households were considered to have either low or medium-low (53%) living standards, but in 2018 the majority of the household was classified as medium-high and high (60%). Across all years most of the sample lived in States with a low or medium level of deprivation. Nevertheless, individuals that live in States with high levels of deprivation increased throughout the years (16% in 1988 vs 24% in 2018), and the proportion of people living in States considered as having very low levels of deprivation decreased (23% in 1988 vs 8% in 2018). Another geographical variable is the region where individuals lived at the time in which the information was collected. The proportions shown in Table 2.3 across years are the same, as this circumstance was also used as a matching variable. Thus, the figures are fixed to the 1988 baseline year, where 38% of the individuals lived in the Central region of Mexico.

Table 2.3 shows the description of the sample according to their *efforts*. Data about the adult's diet show that overall, as individuals aged, the consumption of low nutritious and high-energy food decreased, as well as highly nutritious food. The consumption of maize products and tortillas increased. It is relevant to note that around 6% of the individuals reported the consumption of food supplements in 2006, and the figure increased slightly to 8% in 2018. Figure B.5.1, in the

Table 2.3: Description of Circumstances and Efforts across time

Expected age cohort	0-5		11-16		18-23		24-29		28-33		30-35	
Survey year	**1988**		**1999**		**2006**		**2012**		**2016**		**2018**	
	N	Mean	N	Mean	N	Mean	N	Mean	N	Mean	N	Mean
Circumstances												
Men	6,491	0.51	7,095	0.51	5,660	0.51	9,988	0.51	2,223	0.51	2,208	0.51
Non-Ind	6,491	0.94	4,704	0.92	5,659	0.96	9,988	0.93	2,223	0.89	2,208	0.91
Mother's circumstances												
PNonD	5,683	0.99					1,176	0.73	1,032	0.66	2,018	0.62
BMI_M	5,532	23.85										
M.Anae	3,166	0.12										
HI	6,469	0.49	6,946	0.44								
No Edu	5,669	0.12	6,159	0.16								
Pri.	5,669	0.59	6,159	0.55								
Sec.	5,669	0.21	6,159	0.14								
HS	5,669	0.04	6,159	0.06								
HE	5,669	0.05	6,159	0.00								
Household standards												
WIH	6,471	0.51	5,945	0.41	5,281	0.21	9,988	0.62	2,223	0.55	2,208	0.61
HLS: Low	6,236	0.29	5,208	0.26	5,122	0.22	9,501	0.24	2,216	0.25	2,208	0.18
HLS: MLow	6,236	0.22	5,208	0.26	5,122	0.26	9,501	0.25	2,216	0.24	2,208	0.23
HLS: MHigh	6,236	0.27	5,208	0.26	5,122	0.26	9,501	0.26	2,216	0.24	2,208	0.26
HLS: High	6,236	0.22	5,208	0.22	5,122	0.26	9,501	0.24	2,216	0.26	2,208	0.33
State deprivation												
SDL: V.High	6,491	0.21	7,095	0.15	5,660	0.11	9,988	0.12	2,223	0.09	2,208	0.12
SDL: High	6,491	0.16	7,095	0.22	5,660	0.24	9,988	0.23	2,223	0.25	2,208	0.24
SDL: Med	6,491	0.04	7,095	0.11	5,660	0.21	9,988	0.27	2,223	0.21	2,208	0.32
SDL: Low	6,491	0.35	7,095	0.33	5,660	0.30	9,988	0.23	2,223	0.30	2,208	0.24
SDL: V.Low	6,491	0.23	7,095	0.18	5,660	0.14	9,988	0.16	2,223	0.15	2,208	0.08
Geographic regions												
NW	6,491	0.07	7,095	0.07	5,660	0.07	9,988	0.07	2,223	0.07	2,208	0.07
NE	6,491	0.24	7,095	0.24	5,660	0.24	9,988	0.24	2,223	0.24	2,208	0.24
W	6,491	0.14	7,095	0.14	5,660	0.14	9,988	0.14	2,223	0.14	2,208	0.14
C	6,491	0.37	7,095	0.37	5,660	0.37	9,988	0.37	2,223	0.37	2,208	0.37
S	6,491	0.15	7,095	0.15	5,660	0.15	9,988	0.15	2,223	0.15	2,208	0.15
SE	6,491	0.03	7,095	0.03	5,660	0.03	9,988	0.03	2,223	0.03	2,208	0.03
Efforts												
DP1					2,398	23.13	239	19.65			1,651	18.03
DP2					2,398	20.97	239	18.72			1,651	17.37
DP3					2,398	32.03	239	30.57			1,651	39.07
FS					2,398	0.07	239	0.07			1,651	0.09
PA					5,660	4.26	1,300	6.49			2,197	2.99
Alc					5,660	0.31	289	0.19			2,150	0.19
Tob					5,659	0.25	563	0.48			2,192	0.22

Notes: Matching weights used. Non-Ind=Non-Indigenous; BMI_M=BMI of the mother
M.Anae=Mother anaemic; HI=health insurance; Pri.=Eduaction up to primary school
Sec.=Education up to secondary school; HS=Education up to high school; HE=higher education
PNonD=parents non-diabetic; WIH=water inside the household; HLS=Household living standards,
MLow=Medium-low V.High=Vey high; SDL=State deprivation level, V.Low=Very low; NW=Northwest;
NE=Northeast; W=West; C=Centre; S=South; SE=Southeast
DP1=Low-nutritious, high-energy food (whole diary, fast food, sweetened beverages, candy)
DP2 High-nutritious food (Fruits, vegetables, poultry, fish, cereals)
DP3=Legumes and maize-based products; FS=1 if the individual consumes food supplements
PA=Daily hours dedicated to vigorous and moderate physical activity
Alc=1 if the person consumes alcohol above recommendation; Tob=1 if the individual smokes tobacco frequently
Variables in blue were used for matching

Appendix also shows the distribution of eating patterns across survey years. In terms of physical activity, the table shows that as individuals got older, the average number of hours dedicated to vigorous and moderate physical activities overall decreased. The same pattern follows alcohol-drinking behaviours, when the cohort was between 18 to 23 years old, the portion of people who drank above the recommendations was 31%, but as those adults aged, this portion diminished to 19% in the survey year 2018, when people were between 30 and 35 years old. Finally, it shows that the portion of individuals reported to be regular smokers decreased, in general, 25% in 2006 and 22% in 2018.

2.5.3 Regression models

By estimating Equation (2.8), under the *ex-ante* approach, we find that the vector of common circumstances explains most of the variation for stunting and the double burden. Overall, non-indigenous individuals have a lower probability of experiencing undernutrition, but a higher likelihood of excess weight. Living in the South or Southern regions, as well as in households in deprived conditions increases the probability of under nutritional health problems and developing over nutritional issues. There are differences in the direction of coefficients over time and we could not identify a clear and consistent pattern in the results. When the cohort is between 18 and 23 years of age, living in households with poor conditions and in Northern States increases the probability to be malnourished (e.g having anaemia and excess adiposity or weight). However, when the cohort is in their middle twenties, deprivation and living in Northern States exhibit a negative relationship with malnutrition and anaemia.

The two-step estimation for *ex-post* IOp shows that the total variance explained by our set of circumstances is greater in dietary patterns and smoking frequency. The most statistically relevant circumstances are individual's sex and household living conditions. The second stage shows the inclusion of circumstances as well as true levels of effort. Taken together, these variables explain no more than 18% of the explained variance of the outcomes, with greater relevance in explaining variation in DBM and HB outcomes. Results from the first and second stages bring evidence about the potential health inequalities transmission channels of direct and indirect effects. For example, if circumstances are statistically significant in the first stage, but not in the second, as with geographical region, suggests that those circumstances have an indirect effect on malnutrition that operates only through the effort channel, dietary patterns in our case Fajardo-Gonzalez (2016). All regression models are displayed in the Appendix, Tables B.6.1 to B.7.4.

2.5.4 Inequality of opportunity: *Ex-ante* approach

Table 2.4 shows the results from the mean-based *ex-ante* analysis for different outcomes when using a set of common circumstances. It is worth noting that results from 1988 are not strictly comparable to the rest of the years, given the different outcomes used. The table depicts the level of IOp in children's outcomes and haemoglobin, BMI and WC for adolescents and adults at the clinical cut-off points for anaemia, excess weight, and excess adiposity. For all cross-sections, IOp was also estimated for an outcome that explicitly accounts for the double burden of malnutrition at the individual level.

Table 2.4: Absolute *Ex-Ante* IOp in outcomes defined according to clinical thresholds

Survey year	**1988**		**1999**	**2006**	**2012**	**2016**	**2018**
Expected age of cohort	0-5		11-16	18-23	24-29	28-33	30-35
Stunting	0.284***						
N	5,766						
Wasting	0.180***						
N	5,839						
UWeight	0.264***						
N	5,984						
		Anaem.	0.212***	0.263***	0.274***	-	0.287***
		N	794	4,898	986		2,060
Over weight (BMI)	0.156***	**Excess weight (BMI)**	0.225***	0.053***	0.059**	-	0.026
N	5,827	**N**	1,946	5,103	1,045		2,109
		Excess adiposity (WC)	0.246***	0.121***	0.098***	-	0.057***
		N	1,510	2,647	976		2,046
DBMc	0.287***	**DBMa**	0.313***	0.324***	0.336***	-	0.260***
N	5,478	**N**	589	3,075	483		835

Notes: *p≤ 0.1, **p≤ 0.05, ***p≤ 0.01

Stunting=Height-for-Age below -2 Z-scores; OW(BMIZ)=Body mass index above 2 Z-scores

Wasting=Weight-for-Height below -2 Z-scores

DBM_c defined as HAZ or WHZ below -2 Z-scores and BMI above +2 Z-scores

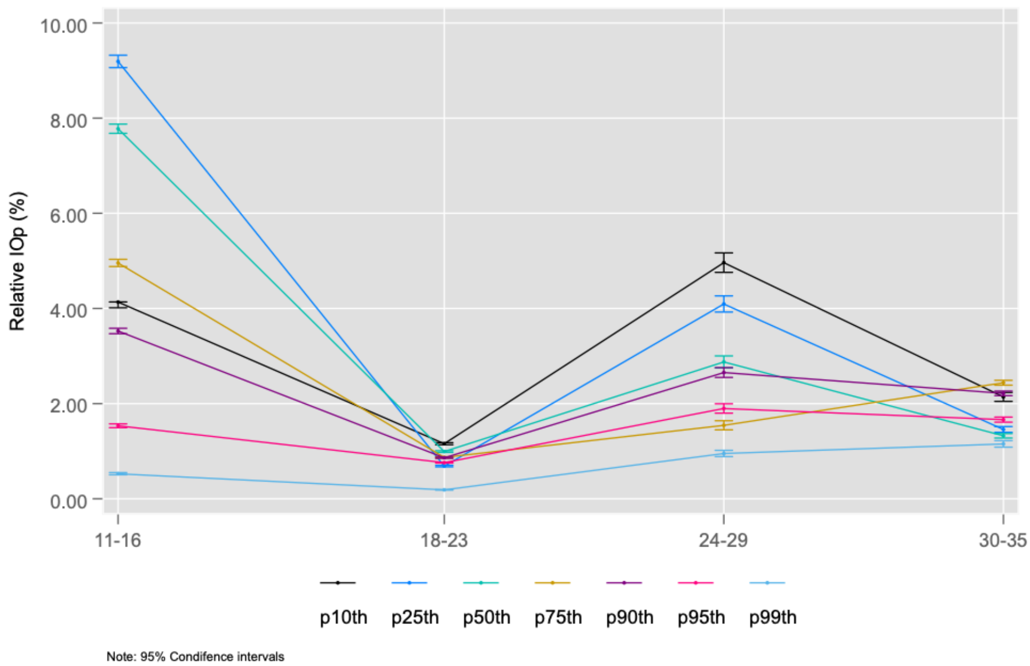
Anaem.=Anaemia=Haemoglobin <13 g/dl; Excess weight (BMI)=Body mass index>25 kg/m^2

Excess adiposity (WC)=Waist circumference> 80cm

DBM_a defined as BMI above 25 kg/m^2 or WC>80 cm and HB<13 g/dl

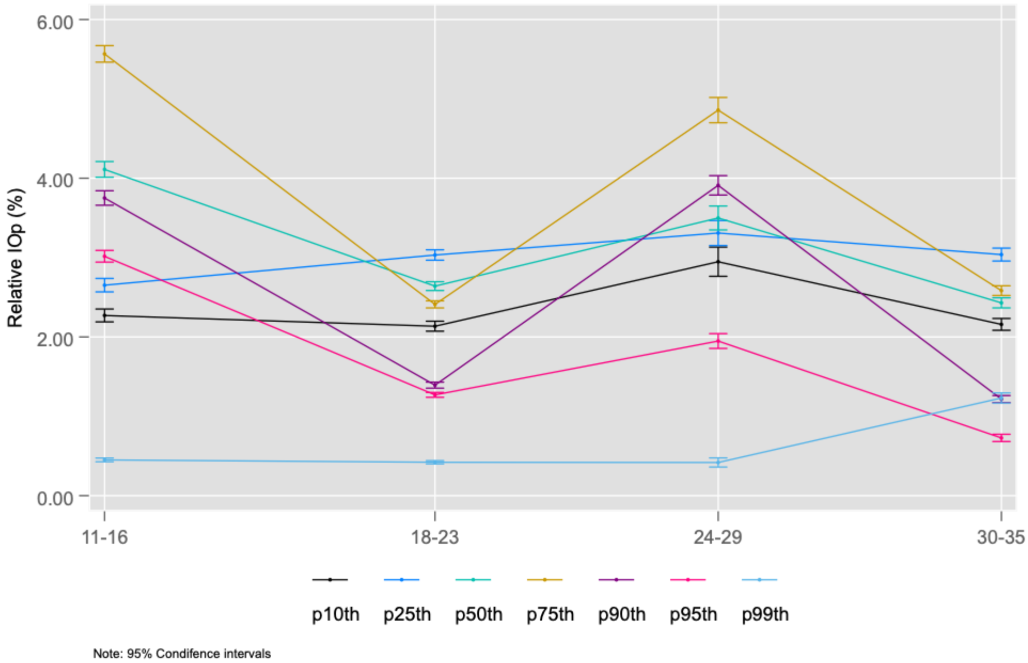
These results show inequality measured through the dissimilarity index, which depicts the fraction of all opportunity sets that should be reallocated from better to worse-off groups to reach an inequality-free situation (Paes de Barros et al., 2008). When the cohort was below five years of age, the level of inequality related to circumstances in stunting, wasting and underweight was 28.4%, 18.0% and 26.4%, respectively, while IOp in overweight was 15.6%. Inequality in child malnutrition was the highest, at 28.7%. Table 2.4 also depicts the level of inequality in anaemia excess weight and central adiposity, these inequality measurements are comparable across years. Overall, the level of IOp in anaemia has increased as the cohort gets older. During adolescence, IOp was 21.2% and when the cohort reached the thirties, inequality reached 28.7%. These results differ for the overnutrition outcomes, where the level of IOp decreased as people aged, from 22.5% to 2.6% in excess weight and from 24.6% to 5.7% in excess adiposity. Thus, in general, inequalities related to circumstances are higher for under compared to overnutrition outcomes. Inequality related to circumstances in malnutrition for adolescents and adults slightly decreased with age, from 31.3% during the cohort's adolescence to 26% in their early adulthood.

Figure 2.2: *Ex-ante* IOp with continuous BMI



Figures 2.2 and 2.3 show the level of *ex-ante* IOp across different points of the BMI and WC distributions. The y-axis depicts the level of relative IOp and the x-axis the ageing of the cohort. Each line is the trajectory of IOp over age for each percentile. The relevance of circumstances in explaining IOp is greater for under than for overnutrition outcomes. For BMI, circumstances matter

Figure 2.3: *Ex-ante* IOp with continuous WC



more for people at the 25^{th} percentile, than for those at the 99^{th} percentile. Inequality related to circumstances across the life course follows a non-linear pattern, in which IOp is higher during adolescence, decreases during emerging adulthood and starts increasing again in early adulthood[13]. Above 18 years of age, circumstances matter more for people at the 10^{th} percentile. Regarding WC, Figure 2.3 shows no clear patterns across percentiles. Overall, the importance of circumstances decreases in emerging adulthood, but increases as adulthood develop up to early adulthood when their importance drops again.

2.5.5 Inequality of opportunity: *Ex-post* approach

In the *ex-post* approach, the total contribution of circumstances and the direct contribution of effort are estimated. Table 2.5 shows the results when estimating Equation (2.17) and computing the relative contribution of circumstances and efforts. To ease the interpretation of the results, circumstances and effort variables were grouped. The table shows the absolute and relative (%) contributions of circumstances and efforts to inequality in anaemia, overweight (either via a proxy of excess weight or excess adiposity) and malnutrition. For anaemia, there is a clear pattern. Across time, circumstances play the most relevant role, around 80% of the total HB variance is explained by circumstances. We did not find concluding evidence about efforts accounting for

[13]Emerging adulthood covers the period between 18 and 29 years of life, while middle adulthood spans from 30 to 45

most of the explained variance in the overnutrition outcomes. On average, efforts contribute to 43% in the variation of excess weight during emerging adulthood, but 54% in middle adulthood. However, circumstances account for most of the variation, 79% and 63% of excess adiposity during emerging and middle adulthood, respectively. Systematically, circumstances are the main driver of inequality in malnutrition, although as people age, the relative importance of circumstances decreases, while that of efforts increases.

Table 2.5: *Ex-post* IOp in outcomes defined according to clinical thresholds for women

Survey year		**2006**		**2012**		**2016**		**2018**	
Expected age cohort		18-23		24-29		28-33		30-35	
Outcome		**Abs.**	**%**	**Abs.**	**%**	**Abs.**	**%**	**Abs.**	**%**
Anaem.	Circum.	0.2780	82.87	-	-	-	-	0.2410	83.97
	Efforts	0.0575	17.13	-	-	-	-	0.0460	16.03
	N		1933	-	-	-	-		1582
EW (BMI)	Circum.	0.0571	56.60	-	-	-	-	0.0219	45.09
	Efforts	0.0438	43.40	-	-	-	-	0.0266	54.91
	N		2033	-	-	-	-		1610
EA (WC)	Circum.	0.1109	79.25	-	-	-	-	0.0444	63.03
	Efforts	0.0290	20.75	-	-	-	-	0.0260	36.97
	N		1064	-	-	-	-		1553
DBM	Circum.	0.3101	76.01	-	-	-	-	0.1968	72.31
	Efforts	0.0978	23.99	-	-	-	-	0.0754	27.69
	N		1227	-	-	-	-		650

Notes: N= observations. Unable to estimate IOp for 2012 and 2016 due to small sample size
Circum.=Total contribution of circumstances. Efforts=Direct contribution of efforts.
Anaem.=Anaemia (HB=Haemoglobin <13 g/dl); EW=Excess weight (BMI=Body mass index > 25kg/m2)
EA=Excess adiposity (WC=Waist circumference > 80 cm)
DBM in adults (BMI > 25kg/m2 or WC>80 cm and HB<13 g/dl)

Figures 2.6 and 2.7 show the results of estimating *ex-post* IOp when the cohort is 10-23 and 30-35 years of age and the contribution of circumstances and efforts is measured at different points of the distributions. This beyond-the-mean analysis shows consistent results regarding the contribution of circumstances and efforts across different parts of the outcome distribution. For BMI, Figure 2.6 shows that across all but the 25^{th} percentile, circumstances are the key contributor to the explained variation. For WC, it was also found that circumstances accounted most for IOp. The figures also display differences in contributions across the lifespan. There are no clear patterns across percentiles and stages of the life span. For both outcome distributions, the only point in common is that circumstances are more relevant for people at the top of the distributions, particularly the percentile 99^{th}. For BMI, efforts are more relevant at the 25^{th} and 50^{th} percentiles, and circumstances become predominantly more relevant than efforts in the 90^{th}, 95^{th} and 99^{th} percentiles. Conversely, for WC, efforts are more relevant at the 90^{th} and 95^{th} percentiles and circumstances have a higher relevance at the 25^{th}, 50^{th}, 75^{th} and 99^{th} percentiles.

2.5.6 Evolution of the contribution of sex, other circumstances and efforts to *ex-post* IOp

Table 2.6 and Figures 2.4 and 2.5 show the results disentangling the contribution of sex, as an independent source of inequalities, circumstances (ethnicity, household living conditions and geographical region) and efforts *ex-post* IOp in all outcomes. Results indicate that sex, as an independent source of inequalities, is of great importance, especially for anaemia (64-58%), excess adiposity (60-32%) and malnutrition (53-48%) in early and young adulthood, respectively. Sex is not relevant for excess weight (0.7-4%). However, there are clear differences across the distribution in the over-nutrition outcomes, Figure 2.4 shows that biological sex is more relevant at the top of the BMI distribution, while Figure 2.5 shows that sex contributes more at the lower percentiles of the WC distribution.

Table 2.6: *Ex-post* IOp in outcomes defined according to clinical thresholds

Survey year		**2006**		**2012**		**2016**		**2018**	
Expected age cohort		18-23		24-29		28-33		30-35	
Outcome		**Abs.**	%	**Abs.**	%	**Abs.**	%	**Abs.**	%
Anaem.	Sex	0.2173	64.78	-	-	-	-	0.1631	56.83
	Circ.*	0.0734	21.89	-	-	-	-	0.0388	13.52
	Efforts	0.0447	13.34	-	-	-	-	0.0851	29.65
	N		1933	-	-	-	-		1582
EW (BMI)	Sex	0.0007	0.70	-	-	-	-	0.0022	4.52
	Circ.*	0.0567	56.21	-	-	-	-	0.0210	43.35
	Efforts	0.0435	43.08	-	-	-	-	0.0253	52.13
	N		2033	-	-	-	-		1610
EA (WC)	Sex	0.0846	60.47	-	-	-	-	0.0227	32.26
	Circ.*	0.0320	22.84	-	-	-	-	0.0213	30.24
	Efforts	0.0234	16.69	-	-	-	-	0.0264	37.50
	N		1064	-	-	-	-		1553
DBM	Sex	0.2187	53.60	-	-	-	-	0.1330	48.86
	Circ.*	0.1145	28.06	-	-	-	-	0.0457	16.78
	Efforts	0.0748	18.34	-	-	-	-	0.0935	34.36
	N		1227	-	-	-	-		650

Notes: N= observations. Unable to estimate IOp for 2012 and 2016 due to small sample size

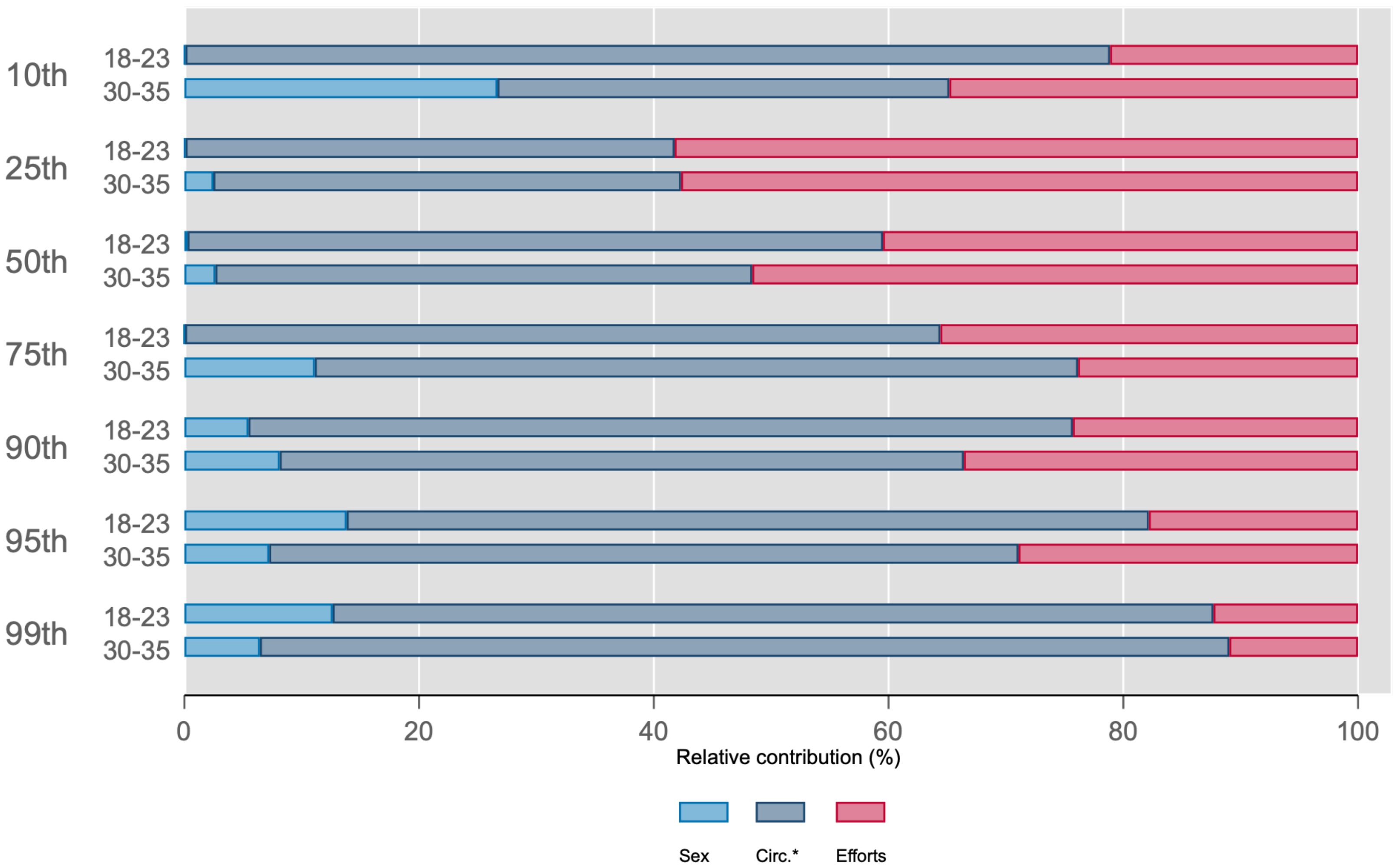
Circ.*=Total contribution of other circumstances excluding sex. Efforts=Direct contribution of efforts.

Anaem.=Anaemia (HB=Haemoglobin < 13 g/dl); EW=Excess weight (BMI=Body mass index > 25kg/m2)

EA=Excess adiposity (WC=Waist circumference > 80 cm)

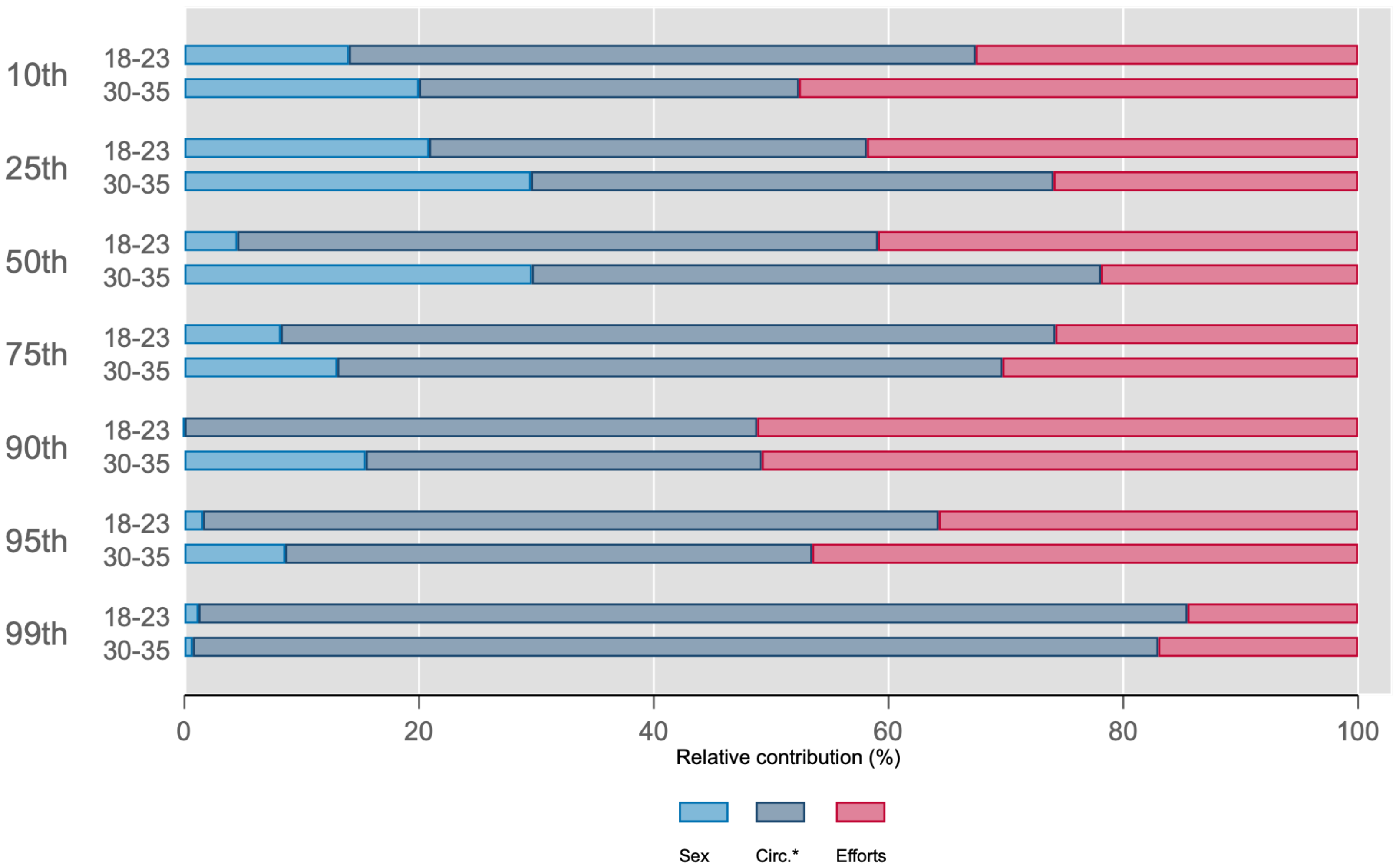
DBM in adults (BMI > 25kg/m2 or WC>80 cm and HB< 13 g/dl)

Figure 2.4: Evolution of the relative contribution of sex, other circumstances and efforts to *ex-post* IOp in BMI



Notes: *Circ. means other circumstances excluding sex. *th* indicates percentile

Figure 2.5: Evolution of the relative contribution of circumstances and efforts to *ex-post* IOp in WC



Notes: *Circ. means other circumstances excluding sex. *th* indicates percentile

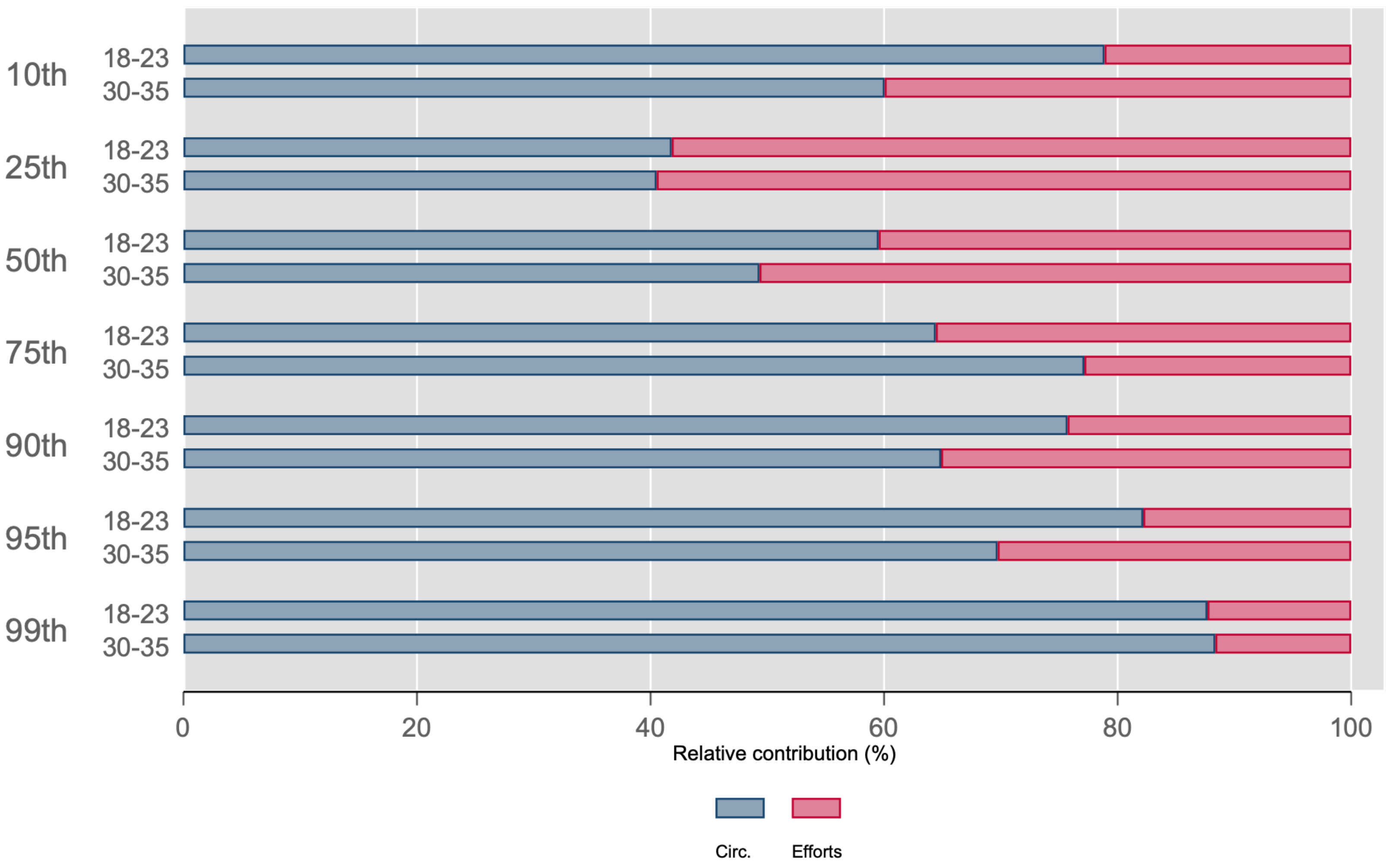
2.5.7 Estimation of *ex-ante* and *ex-post* IOp in malnutrition by sex

The latter results reinforce the importance of conducting a stratified analysis by sex. Tables B.9.1 and B.9.2 display the results from estimating *ex-ante* IOp for both sexes. Overall, girls (0-5 years) observed higher levels of IOp, in all outcomes except underweight, than boys. Although difficult to make comparisons along different periods of the lifespan due to the lack of estimates for boys from 11 to 16 years old, when the cohort was in their early adulthood (18-23 years of age), IOp was higher in men in over and malnutrition outcomes, however, disparities were higher in anaemia. This pattern holds when the cohort is in their young adulthood (30-35), inequalities related to circumstances in undernutrition are higher for women than men, but this latter group observe higher levels of IOp in BMI, WC and malnutrition outcomes. Figures B.9.1 to B.9.4 offer a distributional perspective to *ex-ante* IOp by sex. The results confirm that inequalities in overnutrition tend to be higher in men than women and that inequalities are higher for women sitting at the lower parts of the distributions, while disparities are greater for men at the $10th$, $75th$ and $50th$ percentiles of both distributions.

Tables B.9.3 and B.9.4 display the results when estimating *ex-post* IOp for women and men separately. For men, circumstances are more important than efforts in all outcomes and during both, early and young adulthood. For women, however, there are differences across stages of the lifespan and outcomes. Circumstances are of slightly higher relevance than efforts for undernutrition, but for over and malnutrition circumstances only contribute more than efforts when women were in their early adulthood, as they become older, efforts are the main driver (66%, 62% and 64%) behind inequalities in BMI, WC, and malnutrition, respectively. Overall, this is in line with the prior results about *ex-ante* IOp, in which men observe higher inequalities, compared to women, related exclusively to circumstances. These results reaffirm that efforts matter more for women.

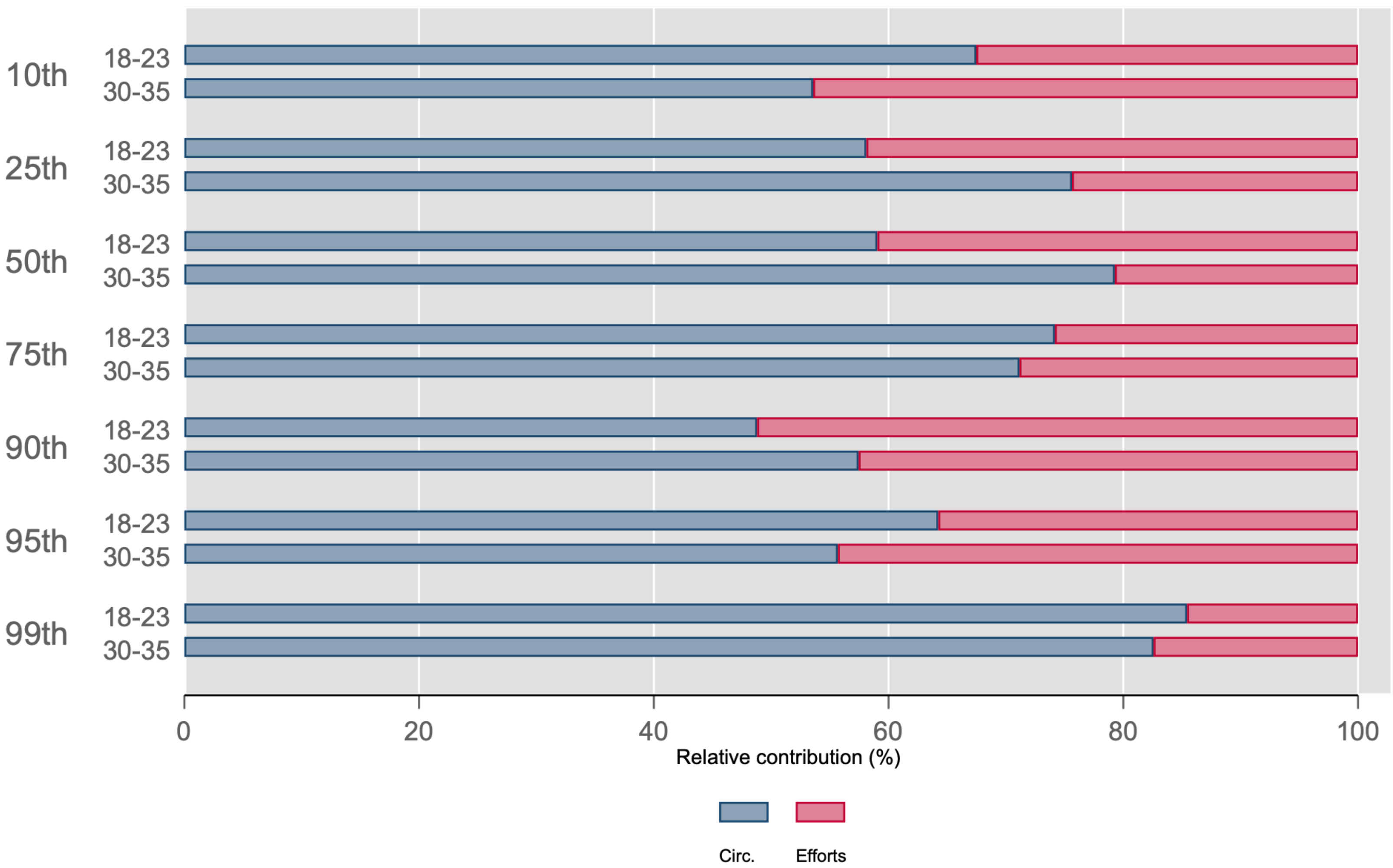
These findings are based on mean-based analyses. Figures B.9.5 to B.9.8 show this circumstances-efforts decomposition across the BMI and WC distribution for women and men. Distinct patterns are observed across outcomes and sexes. Efforts contribute more than circumstances to the explained variation of BMI and WC for women at the bottom ($10th$ and $25th$ percentiles) of the distribution, while there is no clear pattern for men.

Figure 2.6: Evolution of the relative contribution of circumstances and efforts to *ex-post* IOp in BMI



Notes: Circ. means Circumstances. *th* indicates percentile

Figure 2.7: Evolution of the relative contribution of circumstances and efforts to *ex-post* IOp in WC



Notes: Circ. means Circumstances. *th* indicates percentile

2.6 Discussion

This study is a first attempt to deal with the lack of birth cohort data to analyse the evolution of IOp in nutrition-related health outcomes from childhood to adulthood. The study contributes to the ongoing literature by constructing a pseudo-birth cohort for people born between 1983 and 1988. This is as well the first paper to measure IOp in outcomes related to the emerging double burden of malnutrition. The main results of these analyses are the following. First, inequalities that mainly look at the role of circumstances, and therefore are considered unfair, have been persistent across the life course of individuals born between 1983 and 1988. Second, *ex-ante* IOp is higher in undernutrition-related outcomes than in overnutrition. Third, inequalities in anaemia increased as individuals aged, whereas inequalities in excess weight and central adiposity decreased as the cohort got older. Fourth, inequalities related to circumstances in malnutrition have been persistent for 30 years for individuals born between 1983 and 1988. Fifth, when looking at *ex-post* IOp and disentangling the role of circumstances and efforts results indicate that circumstances are the main driver of inequality in anaemia and malnutrition during adulthood. We did not find sufficient and conclusive evidence that efforts are the main driver of variation in outcomes such as excess weight or adiposity.

Despite differences in magnitude, across all outcomes, circumstances drive inequalities. For anaemia, across the lifespan that covers adolescence to early adulthood, between 21% to 28% of the total share of opportunities would need to be reallocated from individuals without anaemia to anaemic individuals to reach equality of opportunity. These numbers are lower for BMI and WC which are 22% during adolescence to 2.6% during emerging adulthood and 24% to 5.7%, respectively. This contrasts with recent evidence from China, where circumstances explained more of the variation of BMI and WC as people aged. Nevertheless, that study was focused on middle-aged and older individuals (Nie et al., 2020). Moreover, the role of circumstances is particularly relevant when looking at the double burden of malnutrition, around 26% to 31% of the share of circumstances would need to be reassigned across the lifespan to reach equality. These results are like those of Fernald et al. (2007) who documented that the prevalence of concurrent overweight or obesity and stunting was approximately 5% in non-indigenous children, and over 10% in indigenous children of two to five years old. This study also found that the factors associated with this double burden were socioeconomic status, maternal age, education, maternal height, and household size (Fernald et al., 2007).

Noticeable, differences between the relative importance of circumstances and efforts to the mean-based level of *ex-post* IOp in excess weight and adiposity outcomes are worthy of discussion. It seems conflicting that while, on average, efforts contribute more to the total variation of excess weight, circumstances are the main driver of *ex-post* inequality in central adiposity and that these patterns are consistent across both points in time. However, the beyond-the-mean results (Figures 2.6 and 2.7) show inconclusive evidence. Although we find consistency at the bottom and upper parts of the distributions, there are mixed results regarding the relevance of circumstances and efforts to IOp in other points of the distribution. For example, although inequality results are similar for people at the 10^{th} and 99^{th} percentiles of both outcome distributions: circumstances contribute more than efforts to the total variation when individuals are 18 to 23 years old. But, when the cohort is older, from 30 to 35 years old, efforts matter the most. At the 90^{th} and 95^{th} percentiles of the BMI distribution, circumstances matter more than efforts, but at the 90^{th} percentile of the WC, efforts are more relevant and at the 99^{th} percentile, circumstances are slightly more important than effort when the cohort is between 18-23 years of age, although efforts play the most relevant role when the cohort is 30 to 35 years old. Thus, this indicates the lack of clear and consistent evidence that efforts are the most relevant factor for inequities in over-nutritional outcomes. At most, we believe that further research is needed to be able to claim that inequality in obesity in Mexico is boosted by people's eating and life-styles patterns. This is an important point to make since there is an implicit belief that most of the variation in over-nutritional outcomes is driven by free will choices. Current policies aiming to tackle the obesity acute crisis in Mexico have framed the roots of obesity and overweight as an individual-decision matter. To be obese is, according to a tacit and collective definition, a deliberate action. Notwithstanding, we did not find enough and conclusive evidence to support this claim.

The additional analyses performed helped to better understand the role that sex, a factor that is clearly and unarguably an exogenous circumstance, has on *ex-post* IOp. Findings showed that, by itself, people's sex account for 58-64% of the explained variation of undernutrition, 32-60% of the explained variance of excess adiposity and 53-48% of the outcome that proxy malnutrition. Individuals' biological sex accounts for very little of the explained variation of BMI (0.7-4%). When splitting the sample by sex and measuring IOp under both approaches, it was found that before exerting any effort and just considering people's circumstances during the first five years of age, female children observed higher levels of inequities in all outcomes, except underweight. However, as individuals reach their young adulthood, inequalities exclusively related to circumstances are higher among male individuals. When incorporating the role of choices and effort-related variables, it was found that these factors are of higher relevance for explaining variations in BMI, WC and malnutrition outcomes among women aged 30-35 years old.

Our results also indicate that inequalities related to circumstances in anaemia display an overall growing trend. Since the role of public programmes is not evaluated in this study, it would be not only interesting but also relevant to explore the effect of the different nutrition-related programmes on the context of IOp under a life-cycle approach. This is pertinent since there is little evidence about the effect of early-life nutritional programmes on later-life outcomes in the case of Mexico. It is of further relevance because, over the last 30 years, there have been a considerable amount of programmes implemented by the Mexican government that has had the clear objective to tackle poverty and undernutrition, the most outstanding example is that of PROGRESA-OPORTUNIDADES-PROSPERA[14], a well-known case study in the policy evaluation literature. As people born between 1983 and 1988 aged, circumstances play a much more relevant role in the level of inequalities related to under-nutritional, and less for over-nutritional outcomes. Paradoxically, most of the nutrition-related programmes in place between 1988 and 2000 aimed to tackle undernutrition in vulnerable people. For instance, the objectives of programmes such as *Liconsa-conasupo*[15], *Tortibonos*[16] or *school breakfasts* to ameliorate the diet of most at-risk people. Impact evaluations of the Liconsa programme found that the fortified milk reduced the risk of anaemia among preschool and school-aged children between 1999 and 2006 (Rivera-Dommarco et al., 2010; Villalpando et al., 2006). Nevertheless, there is yet no evidence regarding whether this effect holds over time, as a lasting protective factor for developing malnutrition later in life. All in all, our analyses show that equality of opportunity in nutrition-related health outcomes is far from being achieved in Mexico. Circumstances are a key source of health disparities for all health outcomes. Therefore, effective policy interventions aimed to enhance people's health should focus more on compensating, rather than on rewarding aspects.

In a more theoretical view, the empirical results of this analysis are also relevant for reconsidering the *economics of obesity* framework. Taking an economic perspective, some studies have claimed that supply-side factors or individual behaviours are at the core of the overweight/obesity epidemic. For example, innovations in food production have resulted in more availability of high-calorie food, and in energy-dense food cheaper than fresh produce (Finkelstein et al., 2010). Others

[14] This is the same programme, although different federal administrations have changed its name

[15] Liconsa is a Mexican parastatal company subsidised by the Federal government. It aims to commercialise fortified milk bags at very low prices for people in extreme poverty and social vulnerability.

[16] *Torti* short for tortilla and *bono* meaning voucher. This programme consisted in granting people in extreme poverty a voucher that could be exchanged for a kilo of tortillas.

have argued that since medical advances have lowered the perceptions of the ill-health effects of obesity, individuals are more likely to be unafraid of being obese (Lakdawalla et al., 2002). Other studies have looked at the role of time preference and obesity, pointing out that people with a high time preference are more willing to enjoy the utility that overeating represents than the future benefits of not doing so (Cavaliere et al., 2013; Zhang et al., 2008). It is as well important to discuss the extent to which these visions are closely related to the idea of a *pure* free will. The IOp framework claims that free will is not entirely orthogonal to circumstances and instead, it inherently depends on individual circumstances. People's decisions are bounded by their structural conditions and available resources, but structural conditions are not given as endowments to individuals, sometimes they are inherited. People do not decide their initial conditions. In this context, the role of government is not only to cope with market failures but also to guarantee the fundamental initial conditions for a healthy start to life.

Grossman's *demand for health* model, which assumes that individuals *inherit* an initial amount of *health* that depreciates with age and increases with investments (Grossman, 1972) is also a case for reflection. Results from this research show that health is not a homogeneous endowment for everyone and that family characteristics shape people's choices later in life and therefore, the rate of health depreciation. This critique is in line with Cunha et al. (2007, 2008, 2009) and Heckman (2012)'s argument in the sense that early environmental conditions are relevant for health outcomes later in an individual's life, although these authors based their argument on the role of cognitive and non-cognitive capabilities. The IOp framework highlights that the amount of the initial stock of health is endogenous to parents' health and their structural conditions. One of the main results of this study is that circumstances explain between 18% to 28% of the variation in nutrition-health outcomes (HAZ, WHZ, WAZ, BMI-Z and DBM) in children. Conceptually, children's circumstances reflect parents' circumstances and their efforts. Mother's health insurance, education, health condition, as well as geographical factors play a non-trivial role in the development of stunning, wasting, low weight for their age and a higher BMI. Moreover, these results show that the rate of health depreciation varies according to the initial amount of inherited health.

It should be emphasised that this is not a causal analysis. We are not interested in measuring the causal effect of circumstances and efforts on nutrition-related health outcomes, but rather to establish the pathways from circumstances to health outcomes and evaluating mediator factors such as individual efforts. As with any analysis, there are some limitations worthy of discussion. One is that we are restricted to the characteristics of the 1988 survey design. For instance, the 1988 survey collected data only about women between 12 and 49 years old, implying that family characteristics (education, BMI, Anaemia) from both parents are missing. Another important point about the data is that the sample units in all the six surveys were the household and hence respondents were those individuals residing in households at the time of the study. This impacted in small sample sizes that did not allow us to measure *ex-post* IOp in 2012 and 2016. We did not estimate *ex-ante* and *ex-post* inequality as well in 2016 because the sampling design was different and data on food consumption were collected differently concerning previous waves, and therefore effort variables were not available. Regarding the data, it is worth adding that food consumption and physical activity information was self-reported. Despite these limitations, the study contributes to different aspects. In terms of the methods, we proposed an innovative way to deal with the lack of panel data when performing life course analysis. With regards to the literature on IOp, this paper is a first attempt to measure the importance of circumstances and efforts to explain a relatively new and certainly worrying public health problem: the double burden of malnutrition. Also, this piece of work contributes to a better understanding of the relevance of studying malnutrition as a two-dimension and interconnected public health problem, instead of assuming that under and overnutrition are independent and exclusive to certain age groups. Future analyses should be undertaken in other settings and using different data to test whether similar results hold.

Chapter 3

Explaining the ethnic gaps in COVID-19 outcomes in Mexico

Abstract. Indigenous people are one of the most socially vulnerable groups across societies. Concerns have been raised about the possibility of greater health disparities when the COVID-19 pandemic interacts with non-communicable diseases in contexts of high socioeconomic inequalities. Using national and administrative public data on COVID-19, this study investigates this hypothesis by explaining differences in COVID-19 health outcomes (hospitalisations and mortality) between indigenous and non-indigenous groups in Mexico. The analysis uses an adaptation of the Oaxaca decomposition method to account for nonlinear responses. This enables the identification and characterisation of the factors behind ethnic disparities. Results indicate that indigenous people have worse COVID-19 health outcomes. These differences are mainly attributable to differences in people's characteristics. Disentangling the contribution of each individual and contextual circumstances to the observable differences, we found that underlying health conditions, and household and municipal socioeconomic characteristics are the main drivers of observable inequalities. These findings highlight that the COVID-19 pandemic exacerbated the pre-existing and longstanding health inequalities between indigenous and non-indigenous people in Mexico.

Keywords: Health inequalities, COVID-19, Oaxaca decomposition; Indigenous groups; Mexico

3.1 Introduction

Higher health disparities could be observed when COVID-19 interacts with a high prevalence of non-communicable diseases (NCDs) among social and economically unequal populations (Horton, 2020). This hypothetical situation could be aggravated if, within societies, a high number of vulnerable groups exist. Based on past experiences, the World Health Organisation (WHO) warned that epidemics have disproportionate effects on vulnerable populations, such as indigenous people, and perpetuate the pre-existing and longstanding social, economic and health inequalities (Sachs et al., 2020). There is evidence that these conditions could be held for the Mexican case. First, before COVID-19, Mexico was already facing a public health crisis mainly driven by non-communicable diseases: 75% of the adult population was either obese or overweight (National Institute of Public Health, 2018) and the type II diabetes prevalence was one of the highest globally at around 13-22%, compared with 6.5%, on average, among OECD (Organisation for Economic Co-operation and Development) countries (Meza et al., 2015; OECD, 2011; Saeedi et al., 2019). Second, Latin America is one of the most unequal regions in the world, and Mexico is not the exception; economic inequality, measured via the Gini index, is one of the highest globally, at 0.45 in 2016 (Lambert et al., 2019). Third, Mexico is a multi-ethnic country. According to the National Institute of Statistics and Geography (INEGI), 21.5% of the population self-identify as indigenous (INEGI, 2020). Indeed, indigenous populations are found all across the country[1].

The unequal impacts of an epidemic on vulnerable populations are closely related to the social and economic circumstances, as well as people's health conditions since such circumstances and conditions influence a wide range of health risks and outcomes (Tai et al., 2020, p. 2). Indigenous people are particularly vulnerable to the COVID-19 pandemic due to the impoverished social and economic characteristics they face. The National Council for the Evaluation of Social Development Policy in Mexico (CONEVAL, in Spanish) estimated that, in 2016, 15.1% of indigenous people did not have access to health services and 56.3% did not inhabit a household with basic standards, such as proper walls, roofs, floors, available running water, a toilet, drainage system or electricity (National Council for the Evaluation of Social Development Policy, 2018). Furthermore, disparities between indigenous and non-indigenous people also exist in other spheres: indigenous people face poorer academic performance, higher levels of poverty, lower life expectancy and health insurance coverage across indigenous populations is still insufficient (Leyva-Flores et al., 2014; Leyva-Flores et al., 2013; Servan-Mori et al., 2014).

[1]Refer to Figure C.0.1 in the Appendix.

Given this context, this study investigates whether the hypothesis regarding the perpetuation of health disparities among indigenous and non-indigenous people holds in light of the COVID-19 pandemic and, if so, the extent to which inequalities have widened. Previous studies have already analysed ethnic inequalities in COVID-19 outcomes, such as hospitalisation and mortality in Mexico. These studies have relied on the estimation of predicted probabilities to observe an outcome, conditional on a set of individual, social and economic characteristics and compare these estimations across groups. Findings from these studies highlight that differences in access and the quality of care have played a crucial role in higher mortality rates among indigenous people compared to the general population (Ibarra-Nava et al., 2021). Also, that living in municipalities with high social deprivation is associated with a higher risk of hospitalisation and early death due to COVID-19 and the presence of underlying health conditions increases the probability of hospitalisation and death among indigenous patients (Serván-Mori et al., 2021). There has been a boom in the literature about inequality and COVID-19. However, a relatively small number of papers have focused on ethnic inequalities. Among these, Yashadhana et al. (2020) studied the case of observing a higher risk of COVID-19 given pre-existing socioeconomic inequalities between the general population and indigenous Australians. In Latin America, Soares et al. (2021) analysed inequalities in excess mortality between indigenous and non-indigenous people in Brazil. For the case of Mexico, Serván-Mori et al. (2021) investigated the variation in hospitalisation and deaths between indigenous and non-indigenous people with COVID-19. This research, in particular, focuses on the study of inequalities in COVID-19 complications and to what extent these are due to people's underlying health conditions and socioeconomic characteristics. The two main health outcomes studied, hospitalisations and deaths due to COVID-19, are meant to be interpreted as "bad" outcomes that reflected a worsening health condition related to COVID-19 infection.

This study contributes to the current literature about ethnic inequalities and COVID-19 by analysing inequalities using an appropriate method to do so. This study goes further than previous studies by decomposing ethnic disparities and investigating potential discriminating effects against indigenous populations. This piece of research unveils discrepancies between high-level commitments to prioritise indigenous people and current social and health policies in practice by identifying ethnic disparities, and their main contributors, which are of relevance for policy targeting purposes. To follow our aim, we first identify ethnic gaps in COVID-19 outcomes and then breakdown this gap into two components, one that explains differences due to observed characteristics and another that explains them given differences in the link between characteristics and outcomes, the latter differences could be attributable to discrimination towards indigenous people. We focus on potential ethnic inequalities that took place before the vaccination programme in Mexico for the general population, which started in March 2021. The remaining part of the paper proceeds as follows: Section 2 explains the decomposition strategy. Section 3 describes the data, while Section 4 presents the key variables used for the empirical analysis. Section 5 reports the main results of the analysis and the last section discusses them.

3.2 Methods

To investigate the extent of COVID-19 outcome differences between indigenous and non-indigenous people, we make use of the nonlinear version of the Oaxaca-Blinder decomposition method, which estimates the impact of individual and contextual characteristics for each group on outcomes and decomposes the average inter-group gap due to differences in *observable characteristics* and differences in *the effects of coefficients*. The standard Oaxaca-Blinder decomposition is based on a regression model where a health outcome is a function of a set of covariates, which in this analysis are: individual's health conditions, household deprivation, health infrastructure individuals are exposed to and the geographical economic characteristics of where people live. This model is run separately for indigenous and non-indigenous people.

To decompose the average inter-group difference, the method relies on a counterfactual that depicts what would happen if the characteristics of one group were interchanged with the coefficients of the other group. By applying this counterfactual, two components are obtained: the *explained* and *unexplained* components. This is known as the *two-fold* Oaxaca decomposition. The former component shows a counterfactual comparison of the expected difference in outcomes if non-indigenous were given the indigenous distribution of covariates. It is explained because this part of the difference can be attributable to differences in the *observed characteristics*. In contrast, the unexplained component reflects a counterfactual comparison of the expected difference if indigenous people experienced the non-indigenous response to the set of covariates, thus it shows the *effect coefficients*. While the explained component might justify group disparities due to differences in people's characteristics, the unexplained part has been labelled as *discrimination*, since there is no economic justification for group differences (Blinder, 1973; Rahimi et al., 2021). As expected, this claim has been controversial, as the concept of discrimination cannot be simply reduced to what a model cannot explain and, at most, it should be labelled as *observed*, incorporating the fact that results are limited to those observable factors included in the model(Jann, 2008).

An additional and relevant point regarding the interpretation of this component, although beyond the scope of this work, is that of Fortin et al. (2011) who explains that the link between the decomposition methods with the impact evaluation literature is that the *unexplained component* of the Oaxaca decomposition can be interpreted as the *population treatment effect on the treated* (PATT) if selection on observables is assumed and holds for identifying treatment effects. This vision has also been shared by Słoczyński (2015) and Słoczyński (2020). Jann (2008) also mentions that the unexplained part captures all the potential effects of differences in *unobserved* variables. In what follows, the linear and nonlinear versions of the decomposition method are described.

3.2.1 Linear model

Aggregate decomposition

Formally, the standard Oaxaca-Blinder decomposition starts with the following structural function[2]:

$$Y_i^g = m^g(X_i, \epsilon_i) = \beta_0^g + \beta_1^g X_{1i} + ... + \beta_k^g X_{ki} + \epsilon_i^g, \qquad g = 0, 1 \tag{3.1}$$

Where Y^g represents the health outcome for group g. X_k depicts different factors that influence the outcome, i indexes individuals and g represents the comparison and reference groups and ϵ_i^g is the idiosyncratic error term of the model. It assumes *additive linearity*: $m(X, \epsilon) = X\beta^g + \epsilon^g$. This implies that the effect of observed and unobserved characteristics are additively separable in $m()$, and it further assumes *zero conditional mean independence*, $E(\epsilon \mid X, G) = 0$.
Thus, the average group difference can be expressed as:

$$\begin{aligned}
\Delta^\mu =& \mu(F_{Y|G=0}) - \mu(F_{Y|G=1}) = E(Y^0 \mid G = 0) - E(Y^1 \mid G = 1) \\
=& E(X\beta^0 + \epsilon \mid G = 0) - E(X\beta^1 + \epsilon \mid G = 1) \\
=& (E(X\beta^0 \mid G = 0) + E(\epsilon \mid G = 0)) - (E(X\beta^1 \mid G = 1) + E(\epsilon \mid G = 1)) \\
=& E(X\beta^0 \mid G = 0) - E(X\beta^1 \mid G = 1)
\end{aligned} \tag{3.2}$$

$$\Delta^\mu = E(X \mid G = 0)\beta^0 - E(X \mid G = 1)\beta^1 \tag{3.3}$$

The Oaxaca-Blinder decomposition requires a counterfactual that illustrates what would happen if the characteristics of one group were interchanged with the coefficients of the other group. Thus this counterfactual could be $F_Y^0 \mid G = 1$, which depicts the average expected outcome for group 1

[2]Following Jann (2018)'s notation

if they had the characteristics of group 0.

$$\begin{aligned}\mu(F_{Y^0} \mid G=1) =& E(X\beta^0 + \epsilon \mid G=1)\\ =& E(X\beta^0 \mid G=1)\\ =& E(X \mid G=1)\beta^0\end{aligned} \tag{3.4}$$

Subtracting and adding $E(X \mid G=1)\beta^0$ in Equation (3.3), we obtain:

$$\begin{aligned}\Delta^\mu =& E(X \mid G=0)\beta^0 - E(X \mid G=1)\beta^1\\ =& E(X \mid G=0)\beta^0 - E(X \mid G=1)\beta^0 + E(X \mid G=1)\beta^0 - E(X \mid G=1)\beta^1\\ =& (E(X \mid G=0) - E(X \mid G=1))\beta^0 + E(X \mid G=1)(\beta^0 - \beta^1)\\ \Delta^\mu =& \Delta^\mu_X + \Delta^\mu_\beta\end{aligned} \tag{3.5}$$

β^g can be estimated using linear regression on the $G=g$ sub-sample and $E(X \mid G=g)$ is the vector of means of X in the same sub-sample. If $\hat{\beta}^g$ is the estimate of β^g and $\bar{X}^g = \hat{E}(X \mid G=g)$ of $E(X \mid G=g)$, the decomposition estimate can be written as follows:

$$\hat{\Delta}^\mu = \hat{\Delta}^\mu_X + \hat{\Delta}^\mu_\beta = \underbrace{(\bar{X}^0 - \bar{X}^1)\hat{\beta}^0}_{\text{Explained}} + \underbrace{\bar{X}^1(\hat{\beta}^0 - \hat{\beta}^1)}_{\text{Unexplained}} \tag{3.6}$$

The decomposition depicted in 3.6 is seen from group 1's perspective, as this is taken as the reference group. If this was changed, the results of the decomposition would change. This issue is known as the *indexing problem* (Cotton, 1988; Neumark, 1988) and implies that results are not unique and depend on the group chosen as reference. The decision of which group to take as a reference should be made based on a preconception of discrimination if this exists. In our case, given the consistent evidence about the unequal treatment between indigenous and non-indigenous people in Mexico (Leyva-Flores et al., 2014; Leyva-Flores et al., 2013; National Council for the Evaluation of Social Development Policy, 2018; Servan-Mori et al., 2014), we believe that the assumption of discrimination against indigenous people holds and therefore, we undertake the decompositions using indigenous people as the reference group.

An additional issue of the Oaxaca-Blinder detailed decomposition, irrespective of the type of model used, is the *identification problem* (Fortin et al., 2011; Oaxaca et al., 1999; Yun, 2005), which refers to the fact that the contribution of binary variables to the unexplained component is not invariant to the choice of the reference category. There is no problem with the explained component, as the sum of the contributions does not change with a change in the reference category. The identification problem is an issue for the detailed decomposition of the unexplained component as it is not possible to disentangle the contribution to the difference in intercepts or to differences in beta coefficients. Thus, the choice of the reference category changes the results (Jann, 2008; Yun, 2005). In this regard, we follow the solution proposed by Yun (2005) which consists of averaging across n sets of estimates produced by varying the reference group across a categorical variable. This is equivalent to estimating a *normalised*[3] equation that identifies the intercept and coefficients of the n dummy variables and averaging the estimates obtained by permuting the reference groups and then using these estimates to perform the decomposition analysis.

Detailed decomposition

The two-fold decomposition previously described is known as **the aggregate Oaxaca decomposition**. But, for many purposes, policy-related included, it is relevant to further identify and measure the main factors contributing to the explained and unexplained parts of the ethnic gap. This extension of the aggregate decomposition is known as **the detailed Oaxaca-Blinder decomposition** and consists of subdividing each component and estimating the contribution of each explanatory variable (Fortin et al., 2011). Thus, given the assumption of *additive linearity*, both the explained and unexplained parts can be further decomposed to disentangle the contribution of the k^{th} explanatory variable to the ethnic gap. Thus, from Equation (3.6) the explained part can be decomposed as:

$$\begin{aligned}\hat{\Delta}^{\mu}_{X} =&(\bar{X}^0 - \bar{X}^1)\hat{\beta}^0 \\ =&\sum_{k=1}^{K}\hat{\beta}^0_k(\bar{X}^0_k - \bar{X}^1_k) \\ =&\hat{\beta}^0_1(\bar{X}^0_1 - \bar{X}^1_1) + \hat{\beta}^0_2(\bar{X}^0_2 - \bar{X}^1_2) + ... + \hat{\beta}^0_k(\bar{X}^0_k - \bar{X}^1_k)\end{aligned} \tag{3.7}$$

and the unexplained part can be decomposed as:

$$\begin{aligned}\hat{\Delta}^{\mu}_{\beta} =&\bar{X}^1(\hat{\beta}^0 - \hat{\beta}^1) \\ =&\underbrace{(\hat{\beta_0}^0 - \hat{\beta_0}^1)}_{\text{Intercepts}} + \sum_{k=1}^{K}(\hat{\beta_k}^0 - \hat{\beta_k}^1)\bar{X}^1_k \\ =&(\hat{\beta_0}^0 - \hat{\beta_0}^1) + (\hat{\beta_1}^0 - \hat{\beta_1}^1)\bar{X}^1_1 + (\hat{\beta_2}^0 - \hat{\beta_2}^1)\bar{X}^1_2 + ... + (\hat{\beta_k}^0 - \hat{\beta_k}^1)\bar{X}^1_k\end{aligned} \tag{3.8}$$

[3] Powers et al. (2000) define normalisation as the process by which one sets the arbitrary constraints to achieve a unique identification of model parameters.

Moreover, within each component, variables of k can be aggregated into subsets, for example:

$$\hat{\Delta}_X^{\mu} = \sum_{k=1}^{a} \hat{\beta}_k^0(\bar{X}_k^0 - \bar{X}_k^1) + \sum_{k=a+1}^{b} \hat{\beta}_k^0(\bar{X}_k^0 - \bar{X}_k^1) + ... \tag{3.9}$$

and for the unexplained part:

$$\hat{\Delta}_\beta^{\mu} = \underbrace{(\hat{\beta_0}^0 - \hat{\beta_0}^1)}_{\text{Intercepts}} + \sum_{k=1}^{a}(\hat{\beta_k}^0 - \hat{\beta_k}^1)\bar{X}_k^1 + \sum_{k=a+1}^{b}(\hat{\beta_k}^0 - \hat{\beta_k}^1)\bar{X}_k^1 + ... \tag{3.10}$$

3.2.2 Nonlinear model

Aggregate decomposition

The set-up is the same, we are interested in an average decomposition. Nevertheless, the first challenge in the nonlinear model is that $E(Y \mid X) = F(X\beta) \neq F(\overline{X}\beta)$. With the nonlinear case it is not possible to insert $E(X)$ into $F(.)$ to get $E(Y)$. Therefore, the difficulty is to generate $\hat{Y} = \hat{E}(Y \mid X) = F(X; \hat{\beta})$ and this implies knowing the functional form for $F(.)$. In this respect, Yun (2004) states that any aggregate Oaxaca-Blinder decomposition is feasible as long as the function $F(.)$ is once-differentiable. Fairlie (1999) for example, proposed an extension of the Oaxaca decomposition using the logit function. According to Fairlie (2005), the decomposition of a nonlinear equation such as $Y = F(X; \hat{\beta})$ can also be written as:

$$\hat{\Delta}^{\mu} = \underbrace{\left[\frac{1}{N^0}\sum_{G_i=0} F(X_i^0\hat{\beta}^0) - \frac{1}{N^1}\sum_{G_i=1} F(X_i^1\hat{\beta}^0)\right]}_{\text{Explained}} + \underbrace{\left[\frac{1}{N^1}\sum_{G_i=1} F(X^1\hat{\beta}^0) - \frac{1}{N^1}\sum_{G_i=1} F(X^1\hat{\beta}^1)\right]}_{\text{Unexplained}} \tag{3.11}$$

Applied to the logit function, $\hat{\Delta}^{\mu}$ denotes the predicted average difference in the coefficients of the binary outcome of interest and $F(.)$ represents the cumulative distribution function from the logistic distribution: $\frac{1}{1+e^{-X\beta}}$. Fairlie (2005) also points out the useful property of the logit regression in that by including a constant term, the average of the predicted probabilities must equal the proportion of the sample. Equation (3.11) shows that the difference in the predicted average observed outcomes can be decomposed into the explained and unexplained components.

Detailed decomposition

One issue that the nonlinear case faces when estimating a detailed decomposition is the *path dependence problem* (Fortin et al., 2011; Powers et al., 2011; Yun, 2004). Although there are different ways to tackle this problem, we follow the solution proposed by Yun (2004) which is simple but robust: a linearisation around $E(X)\beta$ using a set of weights from a first-order Taylor linearisation around Equation (3.11). This allows us to get the contribution of the covariates to Δ_X^μ and Δ_β^μ as relative contributions fixed at the level of the linear predictor(Jann, 2018). For this, let $\hat{E}(X|G=g)=\bar{X}^g$ and $\hat{E}(F(X\beta)|G=g)=\overline{F(X\beta)}^g$. Thus, the aggregate decomposition can be expressed as:

$$\hat{\Delta}^\mu = \left\{\overline{F(X\hat{\beta}^0)}^0 - \overline{F(X\hat{\beta}^0)}^1\right\} + \left\{\overline{F(X\hat{\beta}^0)}^1 - \overline{F(X\hat{\beta}^1)}^1\right\} = \hat{\Delta}_X^\mu + \hat{\Delta}_\beta^\mu \tag{3.12}$$

The individual contribution of each covariate to the characteristics and coefficients effects can be estimated as (Jann, 2018):

$$\hat{\Delta}^\mu_{X,X_k} = \frac{(\bar{X}_k^0 - \bar{X}_k^1)\hat{\beta}_k^0}{(\bar{X}^0 - \bar{X}^1)\hat{\beta}^0}\hat{\Delta}_X^\mu \tag{3.13}$$

and

$$\hat{\Delta}^\mu_{\beta,\beta_k} = \frac{\bar{X}_k^1(\hat{\beta}_k^0 - \hat{\beta}_k^1)}{\bar{X}^1(\hat{\beta}^0 - \hat{\beta}^1)}\hat{\Delta}_\beta^\mu \tag{3.14}$$

such that $\sum_{i=1}^K \hat{\Delta}^\mu_{X,X_k} = \hat{\Delta}^\mu_X$ and $\sum_{i=1}^K \hat{\Delta}^\mu_{\beta,X_k} = \hat{\Delta}^\mu_\beta$. Thus, Yun (2004) proposes to approximate $\hat{\Delta}^\mu$ by first evaluating the function $F(.)$ at the means of the covariates,

$$\hat{\Delta}^\mu \approx \left[F(\bar{X}^0\hat{\beta}^0) - F(\bar{X}^1\hat{\beta}^0)\right] + \left[F(\bar{X}^1\hat{\beta}^0) - F(\bar{X}^1\hat{\beta}^1)\right] \tag{3.15}$$

and then linearising the differences around $\bar{X}^0\hat{\beta}^0$ and $\bar{X}^1\hat{\beta}^1$ using a first order Taylor expansion (Jann, 2018), as follows:

$$\begin{aligned}\hat{\Delta}^\mu &\approx \left[F(\bar{X}^0\hat{\beta}^0) - F(\bar{X}^1\hat{\beta}^0)\right] + \left[F(\bar{X}^1\hat{\beta}^0) - F(\bar{X}^1\hat{\beta}^1)\right] + R_M \\ &\approx \left[(\bar{X}^0 - \bar{X}^1)\hat{\beta}^0\right]\cdot d^0 + \left[\bar{X}^1(\hat{\beta}^0 - \hat{\beta}^1)\right]\cdot d^1 + R_M + R_T\end{aligned} \tag{3.16}$$

where

$$\begin{aligned}R_M = &\left[\overline{F(X\hat{\beta}^0)}^0 - \overline{F(X\hat{\beta}^0)}^1\right] + \left[\overline{F(X\hat{\beta}^0)}^1 - \overline{F(X\hat{\beta}^1)}^1\right] - \\ &\left[F(\bar{X}^0\hat{\beta}^0) - F(\bar{X}^1\hat{\beta}^0)\right] - \left[F(\bar{X}^1\hat{\beta}^0) - F(\bar{X}^1\hat{\beta}^1)\right]\end{aligned} \tag{3.17}$$

and

$$\begin{aligned}R_T = &\left[F(\bar{X}^0\hat{\beta}^0) - F(\bar{X}^1\hat{\beta}^0)\right] + \left[F(\bar{X}^1\hat{\beta}^0) - F(\bar{X}^1\hat{\beta}^1)\right] - \\ &\left[(\bar{X}^0 - \bar{X}^1)\beta^0\cdot d^0\right] - \left[\bar{X}^1(\beta^0 - \beta^1)\cdot d^1\right]\end{aligned} \tag{3.18}$$

where d^g represents the first derivative of $F(\bar{X}^g\hat{\beta}^g) = \frac{\partial F(\bar{X}^g\hat{\beta}^g)}{\partial(\bar{X}^g\hat{\beta}^g)}$. Yun (2004) also mentions that

R_M and R_T are approximation residuals from the evaluation of the function $F(.)$ at the means values and the linearisation. After this, the set of weights for the explained part can be calculated as:

$$W_{\Delta_{X k}} = \frac{((\bar{X}_k^0 - \bar{X}_k^1)\hat{\beta}_k^0)d^0}{((\bar{X}^0 - \bar{X}^1)\hat{\beta}^0)d^0} = \frac{(\bar{X}_k^0 - \bar{X}_k^1)\hat{\beta}_k^0}{(\bar{X}^0 - \bar{X}^1)\hat{\beta}^0} \tag{3.19}$$

and for the unexplained part as:

$$W_{\Delta_{\beta k}} = \frac{((\hat{\beta}_k^0 - \hat{\beta}_k^1)\bar{X}_k^1)d^1}{((\hat{\beta}^0 - \hat{\beta}^1)\bar{X}^1)d^1} = \frac{(\hat{\beta}_k^0 - \hat{\beta}_k^1)\bar{X}_k^1}{(\hat{\beta}^0 - \hat{\beta}^1)\bar{X}^1} \tag{3.20}$$

and

$$W_{\Delta_{X k}} = W_{\Delta_{\beta k}} = 1 \tag{3.21}$$

The weights, $W_{\Delta_{X k}}$, show the contribution of the k^{th} variable to the linearisation of the explained part according to the magnitude of the mean group difference and accounting for the reference group's effect (Powers et al., 2011). Thus, this detailed decomposition using weights is path invariant. The decomposition can be expressed in terms of the overall components as a sum of weighted sums of the unique contributions, as:

$$\overline{Y}_A - \overline{Y}_B = E + U = \sum_{k=1}^{K} W_{\Delta_{X k}} E + \sum_{k=1}^{K} W_{\Delta_{\beta k}} U = \sum_{k=1}^{K} E_k + \sum_{k=1}^{K} U_k \tag{3.22}$$

Jann (2018) warns that if the volume of data is in highly nonlinear regions of $F(.)$, or differences in coefficients or means are large, the approximation could be poor.

The Oaxaca-Blinder decomposition is possible if average group differences in outcomes exist. So, we examine this by using tests on differences in proportions, and then we estimate nonlinear models for each group and outcome. We also make use of Linear Probability Models (LPM) aiming to check the robustness of the results, although only results from the nonlinear models are reported in the main text, linear model results are shown in the Appendix, Tables C.3.1 and C.3.2, as well as results from testing differences in coefficients across groups using a Wald-type test of nonlinear hypotheses for the estimated parameters.

3.3 Data

This analysis uses open administrative data on COVID-19 from the General Directorate of Epidemiology (Dirección General de Epidemiología, (DGE) in Spanish), the 2020 National Census and the General Directorate of Health Information (DGIS, in Spanish).

3.3.1 COVID-19 data

The Mexican government did not follow universal COVID-19 testing and only those with symptoms were eligible for a test. The data collection process (who, where, when and how data is collected) is described and explained and depicted in Figure C.1.1 in the Appendix. The publicly available dataset was updated every day since the pandemic hit Mexico, and some of the variables might have reporting delays (Giannouchos et al., 2020). For this analysis the version released on the fourth of April 2021 is used. Results about testing, hospitalisation and patient follow-up (discharge, or worsening condition where patients are admitted to ICU (intensive care unit) or passing away) are directly uploaded by the diagnostic facility or hospital according to test results.

The dataset contains information about the patient's birthplace, place of residence, age, sex, nationality and ethnicity (whether a patient identifies as an indigenous language speaker). Information about the patient's migratory status is also included, as well as the patient's health institution affiliation and clinical information. This includes the type of medical attention, for an inpatient: admission date, symptom onset date, whether admission to ICU and/or date of death, polymerase chain reaction (PCR) test result (positive, negative, or pending) and if women tested are pregnant. Additional clinical information about underlying conditions such as pneumonia, chronic obstructive lung disease (COPD), asthma, immunosuppression, diabetes, obesity, hypertension, chronic renal and cardiovascular disease, and other comorbidities is also included, as well as whether the patient is a smoker. These are all indicator variables and no further information is provided in the dataset. Since there is no variable available to identify a patient, we matched cases with the same information about demographics and clinical history data and eliminated duplicate observations[4]. Patients with incomplete (pending results) or missing information about testing results and ethnicity were also excluded, as well as non-Mexican patients, resulting in a final sample of 4,797,854 observations.

3.3.2 2020 National Census Data

For our analysis, we also made use of aggregated information at the municipality level including household characteristics, number of people affiliated with health services and number of health facilities across the country in January 2020, number of people that attended school or were literate

[4] A similar approach was followed by Mancilla-Galindo et al. (2020) in their modelling about COVID-19 deaths in Mexico.

as well as labour-market characteristics such as number of people unemployed. These indicators were constructed by getting the percentage of people with the j characteristic living in the m municipality.

3.4 Key variables

3.4.1 Ethnic groups

Ethnicity is a binary classification with the two groups identified according to whether or not individuals speak an indigenous language[5]. It is worth commenting that another definition that we might have used, that is recurrent in the Mexican literature for ethnicity, is whether individuals belong to an *indigenous household*, despite them not speaking an indigenous language. Unfortunately, we could not use this definition since the dataset identified indigenous individuals according to the language they speak[6].

3.4.2 Health variables

Three measures of health outcomes are used to reflect a worsening condition of people who contract COVID-19:

- To be hospitalised due to COVID-19
- To die because of COVID-19-related complications

Outcome variables are binary indicators and take the value of 1 if the event is true and 0 otherwise. For some individuals, the events are conditional on a prior event being true. However, not all dead patients were either hospitalised.

3.4.3 Individual and structural variables

People are vulnerable to epidemics not only because of their particular health conditions (overall health status, comorbidities or the risky health behaviours they adopt) but also because of their social, economic and household circumstances. COVID-19 has made it explicit that individual and contextual factors matter. A recent study showed that comorbilities such as obesity, diabetes, hypertension, coronary heart disease, and heart failure were closely related to severe COVID-19 cases (Hernández-Garduño, 2020). It has also been found that socioeconomic factors or structural conditions play a role in worsening the impact of the pandemic within communities (Hawkins

[5]There are 68 indigenous languages in Mexico.
[6]A modification in the identification of an indigenous individual in the dataset came later. Since October 7, 2020, it was possible to identify ethnicity according to the patient's ethnicity self-identification.

et al., 2020). In particular, a recent study found that poor access to water; language barriers; household characteristics; lack of health insurance; and underlying health conditions such as hypertension, type II diabetes, chronic pulmonary diseases and respiratory tract infections are risk factors hampering the ability of indigenous communities to avoid contracting COVID-19 (Díaz de León-Martínez et al., 2020). Based on this evidence, the key explanatory variables used in this analysis are divided into two categories, as shown in Table 3.1, individual-level characteristics and socioeconomic circumstances, which have been aggregated at the municipal level. Data at the municipality-of-residence level is used since it is an indirect but reliable way to proxy the social and economic aggregated deficiencies that can be correlated with health outcomes. The institution where individuals received medical attention is captured by a vector that contains eight binary variables. One for each institution that composes the Mexican health system. The variables identify the institution where individuals received medical attention if hospitalised or where they died. These variables take the value of 1 if attention was received in the given institution and 0 if otherwise. The inclusion of these variables is based on previous evidence (García-Peña et al., 2022; Sánchez T., 2020) about COVID-19 treatment differences across the public health institutions that compose the Mexican health system. Further description of the individual level variables can be found in Table C.2.1 in the Appendix.

Table 3.1: Individual and contextual variables used in the analysis

Individual-level characteristics	**Socioeconomic circumstances**
-Demographics	-Household deprivation characteristics
-Underlying health conditions	-Health coverage and medical infrastructure
-Risky behaviours	-Economic characteristics
-Institution where individuals received medical attention*	

* This variable is relevant because the Mexican health system is fragmented into several health institutions and a recent analysis found relevant contrasts in Covid-19 mortality rates within public institutions and between the public and private sectors. For example, up to August of 2020, 45% of hospitalised patients in the IMSS died, versus 31% of the patients hospitalised in SSA hospitals and 16% in the private sector (Sánchez T., 2020)

Socioeconomic circumstances refer to those characteristics that people cannot change in the short-run and that occur within a geographical area, in this case, the municipality. Among these, we include the number of households that live in deprived conditions within a municipality (average number of people per household; percentage of households with low-quality-material walls, ceilings, and floors; percentage of households without electricity; percentage of households without running water or toilet, drainage, electricity or all these; percentage of households with fridges, radio, TV, mobile phone or internet); and the level of health coverage, medical infrastructure (percentage of people without health insurance and number of health facilities available in the municipality in January 2020, when the pandemic started), and municipality's economic conditions (percentage of unemployed people and with no formal education)

3.5 Results

3.5.1 Ethnic differences in COVID-19 outcomes and covariates

Table 3.2 shows the size of the sample for each outcome and group, as well as the proportion of people hospitalised and dead. It shows the difference in these proportions. For example, -0.120 (0.129-0.249) indicates that the proportion of hospitalisations was 0.12 greater among indigenous people than non-indigenous people. The negative sign shows that, across all COVID-19 outcomes, indigenous people were affected more than the non-indigenous population. A test of differences in proportions to determine whether these differences across groups are statistically significant was performed and the associated *p-value* is displayed. All differences are statistically significant.

Table 3.2: Ethnic differences in proportions of people hospitalised and dead due to Covid-19 in Mexico

	N_NI	Mean NI	N_I	Mean I	Diff	p-val
People hospitalised	4,797,799	0.129	31,272	0.249	-0.120	0.000
People dead	4,797,799	0.051	31,272	0.098	-0.048	0.000

Notes: Analysis period Jan 2020- March 2021. NI=Non-indigenous. I=Indigenous.
Diff=Raw difference. Two-sided p-value

Table 3.3 shows the mean value of the individual characteristics for the two health outcomes for each group. Overall, there are statistically significant differences across the groups for most of the characteristics and outcomes. Exemptions can be found in some demographics. Across the outcomes, the proportion of people with a comorbidity is larger for indigenous than for non-indigenous people, whereas the non-indigenous group have a higher proportion of NCD than indigenous. A higher proportion of people with obesity and who are smokers is observed in the non-indigenous group. Table 3.3 shows that diabetes, hypertension and obesity have the highest proportion values across the outcomes and these proportions are greater among non-indigenous people. Concerning medical care, across the two outcomes, most of the individuals received care and were treated in hospitals owned by the Mexican Social Security Institute (IMSS) and the Health Ministry Hospitals

(SSA). Most of the hospitalisations among non-indigenous people took place at IMSS hospitals, while SSA hospitals treated indigenous people. This is something expected as SSA hospitals admit most of the people enrolled to "INSABI" (formerly known as the *Seguro Popular*) programme. Among those who died, most of the non-indigenous individuals were treated in IMSS hospitals and most of the indigenous that died received attention in SSA institutions.

Table 3.3: Ethnic differences in individual characteristics for all outcomes

	Hospitalised				**Died**			
	Mean NI	Mean I	Diff.	p-val	Mean NI	Mean I	Diff.	p-val
Demographics								
Age	54.85	54.64	0.22	0.34	63.08	63.73	-0.65	0.02
Female	0.43	0.44	-0.01	0.01	0.38	0.38	0.00	0.72
Comorbidities								
COPD	0.04	0.07	-0.03	0.00	0.05	0.08	-0.03	0.00
Asthma	0.02	0.03	-0.01	0.00	0.02	0.03	-0.02	0.00
Immunosuppression	0.03	0.03	0.00	1.00	0.03	0.02	0.01	0.04
Renal D.	0.06	0.05	0.01	0.00	0.08	0.06	0.02	0.00
Pneumonia	0.58	0.64	-0.05	0.00	0.70	0.78	-0.08	0.00
Other Comorb.	0.06	0.05	0.01	0.00	0.06	0.05	0.01	0.00
NCD								
Diabetes	0.31	0.30	0.00	0.53	0.38	0.36	0.02	0.01
Hypertension	0.36	0.30	0.06	0.00	0.45	0.38	0.07	0.00
Cardiovascular D.	0.05	0.04	0.01	0.02	0.06	0.05	0.01	0.03
Risky Behaviours								
Obesity	0.19	0.18	0.01	0.01	0.21	0.22	-0.01	0.22
Smoker	0.08	0.06	0.02	0.00	0.08	0.07	0.01	0.02
Medical Attention								
Test waiting-time	4.62	4.50	0.12	0.01	5.08	5.00	0.09	0.23
Private	0.03	0.01	0.02	0.00	0.02	0.01	0.01	0.00
IMSS	0.52	0.27	0.25	0.00	0.60	0.30	0.30	0.00
ISSSTE	0.08	0.07	0.01	0.00	0.07	0.07	-0.00	0.66
SSA	0.31	0.60	-0.29	0.00	0.27	0.58	-0.31	0.00
State	0.02	0.01	0.01	0.00	0.02	0.01	0.01	0.00
PEMEX	0.02	0.00	0.01	0.00	0.01	0.00	0.01	0.00
SEDENA	0.02	0.04	-0.02	0.00	0.01	0.04	-0.03	0.00
SEMAR	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01

Notes: Analysis period Jan 2020- March 2021. NI=Non-indigenous. I=Indigenous. Diff=Raw difference. Two-sided p-value

Table 3.4 describes the average values of the structural socioeconomic circumstances of the areas where individuals lived in 2020 at the municipality level. There are significant differences across all variables. Overall, indigenous people lived in municipalities with a lower per cent of urban localities and in municipalities where the percentage of households in poor physical condition (no floor, less spacious homes, with no water, electricity, drainage, motorcycle, home appliances, TV, radio, telephone, computer or internet) is larger than where non-indigenous people lived. With regards to health insurance coverage, the percentage of people not covered is slightly higher in municipalities where non-indigenous people live. The number of health facilities is, on average, lower in municipalities where indigenous people live. The economic characteristics of municipalities are overall better in non-indigenous municipalities. The average percentage of illiteracy is lower, but the percentage of unemployment is slightly higher. Differences in these indicators across groups are statistically significant.

Table 3.4: Ethnic differences in municipal socioeconomic characteristics

	Mean NI	Mean I	Difference	p-val
Percentage of Urban Localities in the municipality	28.11	15.15	12.95	0.00
Household Characteristics				
Average number of people per household	3.49	3.75	-0.27	0.00
Percentage of households with low-quality-material floors	0.50	1.75	-1.25	0.00
Percentage of households with only one sleeping room	8.46	10.41	-1.95	0.00
Percentage of households with only one room	1.46	2.57	-1.11	0.00
Percentage of households without water	0.57	1.48	-0.90	0.00
Percentage of households without electricity	0.10	0.51	-0.41	0.00
Percentage of households with a latrine	0.39	2.52	-2.14	0.00
Percentage of households without drainage	0.44	3.21	-2.77	0.00
Percentage of households without E,W,D*	0.02	0.17	-0.15	0.00
Percentage of households without car or motorcycle	13.17	15.46	-2.30	0.00
Percentage of households without appliances	0.16	1.08	-0.92	0.00
Percentage of households without TVs or Radio	0.83	2.53	-1.70	0.00
Percentage of households without telephone or mobile phone	1.46	4.37	-2.91	0.00
Percentage of households without computer nor internet service	9.12	14.14	-5.01	0.00
Percentage of households without ICT technologies	0.28	1.48	-1.20	0.00
Health Insurance Coverage and Infrastructure				
Percentage of people not affilated to a health institution	25.68	24.09	1.60	0.00
Number of health facilities in January 2020	148.48	81.12	67.36	0.00
Economic Municipal Characteristics				
Percentage of people from 12-14 years that you did not attend school	0.36	0.53	-0.17	0.00
Percentage of people from 8-14 years who can not read nor write	0.21	0.46	-0.25	0.00
Percentage of people above 15 that is illiterate	2.02	6.23	-4.20	0.00
Percentage of people above 15 without schooling	2.51	5.73	-3.22	0.00
Percentage of unemployed people	1.05	0.79	0.26	0.00

Notes: Data from the 2020 National census. *E,W,D=No electricity, water and drainage.
NI=Non-indigenous. I=Indigenous. Diff=Raw difference. Two-sided p-value

3.5.2 Oaxaca-Blinder decomposition

Aggregate Decomposition

Both Tables, 3.2 and 3.5, present the same information regarding the average health outcome for each group, $\overline{Y}^g$, g=*0,1*. The outcomes take binary values, thus both tables show the proportion of indigenous and non-indigenous people hospitalised and dead. However, Table 3.5 also shows the results of the average gap decomposition. The explained component, which depicts the extent to which differences between groups are due to differences in observable characteristics, accounts for most of the average difference, 82% and 88% for hospitalisations and COVID-related deaths. The unexplained component, which measures the extent to which average differences between groups are due to the link between characteristics and outcomes, contributes positively to the ethnic gap in magnitudes of approximately 18% and 12%, respectively. Comparisons of linear and nonlinear models are found in the Appendix, Tables C.4.1 to C.4.3, but, overall there are small differences in the decomposition results.

Table 3.5: Aggregate Oaxaca Decomposition. Nonlinear models

	Hospitalisations	%	Deaths	%
Non Indigenous	0.127***		0.050***	
	(0.00)		(0.00)	
Indigenous	0.245***		0.097***	
	(0.00)		(0.00)	
Mean Difference	-0.118***		-0.047***	
	(0.00)		(0.00)	
Explained	-0.096***	81.709***	-0.041***	87.634***
	(0.00)	(1.06)	(0.00)	(1.84)
Unexplained	-0.021***	18.291***	-0.006***	12.366***
	(0.00)	(1.06)	(0.00)	(1.84)
Observations	4,796,808	4,796,808	4,796,808	4,796,808

Notes: Bootstrapped standard errors in parenthesis (500 replications)
Models fitted using an ANOVA-type normalisation and weights from
a first-order Taylor linearisation. % share of each component to the overall gap.
+ $p<0.1$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$

Detailed Decomposition

Figures 3.1 and 3.2 show the relative contribution of each sub-set of variables to the explained and unexplained components, respectively. Contributions are shown as a percentage of the overall difference. Positive contributions indicate that if the distribution of a characteristic was swapped between indigenous and non-indigenous people a reduction in the ethnic gap would be expected. Likewise, a negative contribution indicates that if the counterfactual is observed, the ethnic gap is expected to increase. For all results, we provide the uncertainty of our estimations. In particular, we report bootstrapped standard errors based on 500 replications with replacement. Detailed estimations are found in Tables C.4.2 to C.4.3 in the Appendix.

From Figure 3.1, it can be seen that hospitalisations: demographics, underlying conditions, risky behaviours, health infrastructure, household characteristics and municipal economic characteristics positively contribute to the ethnic gap, and medical attention contributes negatively. All contributions are statistically significant, except for municipal economic characteristics. For deaths related to COVID-19, demographics, underlying conditions, risky behaviours and municipal economic characteristics have positive contributions, while medical attention, health infrastructure and household characteristics contribute negatively to the gap. All these contributions are statistically significant, except for health infrastructure, household characteristics and municipal economic characteristics.

The presence of underlying conditions is one of the main drivers of the explained ethnic differences. This means that if indigenous were equal to non-indigenous in the distribution of their comorbidities, the ethnic gap in hospitalisations and deaths would be expected to reduce by 46%

and 51%, respectively. For hospitalisations, household characteristics are the second driver of the explained differences between the ethnic groups. If indigenous people had the same household conditions as non-indigenous people, the ethnic gap would decrease by 33%. Individual demographics are important drivers of explained differences in COVID-19 deaths, by shifting the age and sex distributions of non-indigenous to match the indigenous distribution, the difference in deaths between groups would decrease by around 30%.

While a positive effect indicates a reduction in the gap, a negative sign denotes a potential increase in differences between groups. Health institution where individuals received medical attention is a factor that increases the indigenous/non-indigenous differential. If indigenous people were affiliated with the same health institutions to which non-indigenous people are affiliated, the ethnic difference in hospitalisations and deaths due to COVID-19 would increase by 11% and 8%. The estimations of the detailed decomposition of the unexplained component for all outcomes show a lot of uncertainty as the confidence intervals are very large. The set of demographic variables has a negative contribution. This means that if the link between sex and age with these outcomes is the same across groups, the ethnic differences would increase by 19% and 10%, respectively.

Figure 3.1: Detailed Oaxaca decomposition: Explained component

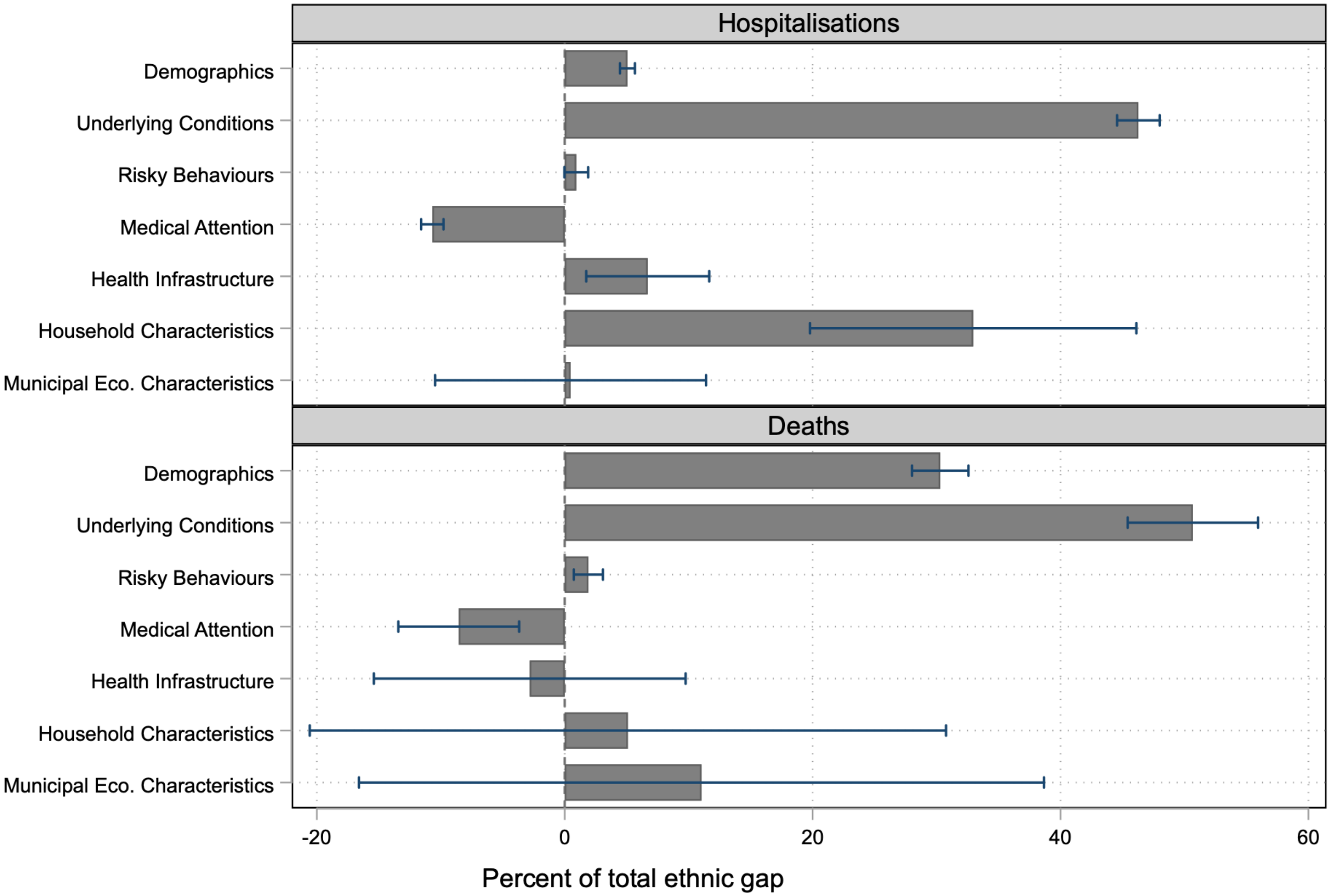
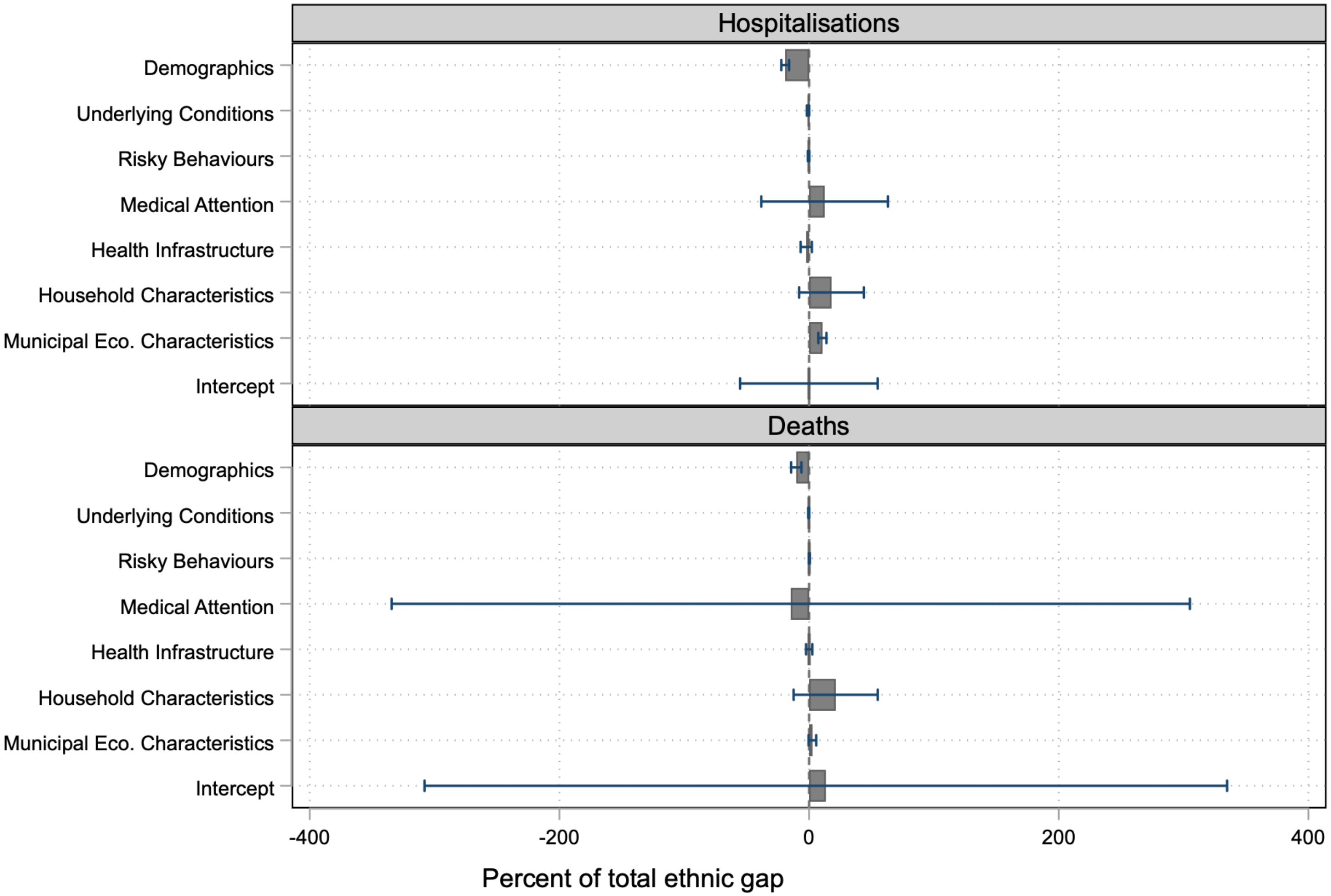


Figure 3.2: Detailed Oaxaca decomposition: Unexplained component



3.6 Discussion

Using administrative data on COVID-19, this analysis identifies and measures the average differences in hospitalisations and deaths due to COVID-19 between indigenous and non-indigenous people in Mexico. This study uses a nonlinear version of the Oaxaca-Blinder decomposition method and finds four main results. First, differences due to individual characteristics account for most of the observed ethnic gap. Once accounting for these characteristics, a non-trivial part of the ethnic gap remains unexplained. Second, people's underlying conditions (comorbidities and non-communicable diseases) are the main driver of the explained differences in hospitalisations and deaths due to COVID-19 between indigenous and non-indigenous people. Third, the health institutions where people received care explain differences in the ethnic gap between hospitalisations and deaths. Fourth, if household conditions were equalised across municipalities, the ethnic gap in hospitalisations due to COVID-19 would decrease by 32%.

Once differences between groups have been explained given observable characteristics, there remains a part that has no economic explanation. This component indicates inter-group differences in the relationship between characteristics and the observed outcomes and has been framed as *observed discrimination*, but this claim needs to be taken with further caution. We argue that although discrimination cannot be defined according to a part that a model cannot explain, this, by itself, does not rule out the possibility that discrimination against indigenous people exists.

The fact that an unequal provision of health and other public services across municipalities exist underpins evidence of a systematic unequal treatment that disadvantages indigenous communities. This is particularly worrying since Mexico signed the *International Labour Organisation (ILO) Indigenous and Tribal Peoples Convention*, better known as ILO Convention #169, and the *Acuerdo de San Andrés* in 1989 and 1996, respectively. Both conventions highlight the aspirations of indigenous people to develop and maintain their identities, languages and religions while exercising their fundamental human rights to the same degree as the rest of the population, and that this must be guaranteed by the State. Thus, in this sense, the two conventions represent a benchmark for policy-making, since it reinforces the rights that indigenous peoples have on top of those they are entitled to by the Mexican Constitution. With regards to health, articles 24 and 25 of the #169 convention state that social security and health services *should be extended* progressively to reach full coverage, that the delivery of health services *should be community-based* and that the health system *should prioritise* the delivery of primary health care services (International Labour

Organisation, 2009). In Australia for example, public policies focused on prevention and primary care have shown to be effective to reduce ethnic health disparities (Davis, 2004; McIntyre et al., 2005). Nevertheless, a study for Mexico showed that indigenous people did not utilise primary care due to the lack of confidence, mistreatment, unavailability and facility's remoteness (Servan-Mori et al., 2014).

Underlying health conditions are a major factor in the explained differences in hospitalisations and deaths due to COVID-19 between indigenous and non-indigenous people. This result is in line with previous studies that pointed out that non-communicable diseases such as diabetes, obesity and hypertension were positively associated with COVID-19 outcomes (Gutierrez et al., 2020; Hernández-Galdamez et al., 2020; Monterrubio-Flores et al., 2020; Serván-Mori et al., 2021). Before the pandemic, Mexico was already facing an acute obesity crisis, a health problem that has not been entirely addressed by the government (Barquera et al., 2020). Thus, our findings highlight that unsolved public health problems make indigenous people more vulnerable. In the light of health shocks, the pre-existing and longstanding health inequalities between indigenous and non-indigenous people magnify, persist and can further perpetuate. This analysis also evidences that disparities in outcomes could potentially come from the Mexican health system itself. The type of health institution where people received medical care for COVID-19 is relevant to explain differences in hospitalisations and deaths due to COVID-19. If medical attention would have been the same between indigenous and non-indigenous people across the health system's institutions, ethnic differences would have increased. This means that differences in the quality of care within the public sphere of the health system exist. This coincides with previous studies, Puig et al. (2009) found high levels of heterogeneity in healthcare quality and that users rated better healthcare attention received in SSA institutions than in IMSS facilities (Puig et al., 2009), it is worth noting that both institutions, SSA and IMSS, belong the public sector. Another study also found profound differences in healthcare quality across institutions that compose the Mexican health system, finding out that the probability to die due to COVID-19 was the highest among people treated in the IMSS (Mexican Social Security Institute). (García-Peña et al., 2022) Sánchez T. (2020) found as well that COVID-19 mortality variation across the institutions of the health system was due to structural differences in the hospital infrastructure, equipment availability and training of the staff, as well as the use of care protocols and that the pandemic only exhibited these deep-rooted inequalities (Sánchez T., 2020). Indeed, Table 3.4 shows that, at the beginning of the pandemic, the number of health facilities was, on average, larger in non-indigenous municipalities. This also highlights the lack of an indigenous-prioritising policy regarding health facilities openings and availability .

Household and municipal socioeconomic conditions matter. This is relevant for contexts where a federal political system prevails. The federal system in Mexico has led to different levels of efficiency, efficacy and quality in the provision of health services across the States that compose the Mexican Federation. Therefore, where people live conditions the services to which they have access. Historically, indigenous settlements have experienced a relatively higher scarcity of health facilities along with low quality of healthcare services (Juárez-Ramírez et al., 2014; Leyva-Flores et al., 2014; Leyva-Flores et al., 2013; Servan-Mori et al., 2014). Furthermore, a study found that living in areas with low healthcare resources was associated with a higher risk of hospitalisation for COVID-19 (Serván-Mori et al., 2021). This highlights the need to consolidate a coordinating and responsive federal system that can guarantee universal health insurance coverage and access to basic medical care for all citizens, regardless of their ethnicity or postcode.

In terms of the methods used, our results corroborate previous conceptions about similarities in results between linear and nonlinear models when the outcome variable is binary, although nonlinear models are better when the gaps are located in the tails of the distribution (Fairlie, 2005). This study is not without limitations. The most challenging is the under-representation of the *real* number of deaths due to this pandemic. Since barriers to access the health system exist, many people died in their homes and therefore were not registered in the administrative dataset we used (Soberanes, 2021). Further analysis is needed to investigate whether this event increased the mortality ethnic gap and who was affected the most. Also worthy of attention in future studies is, for example, extending the cross-sectional analysis and decomposing mean group differences over time.

All in all, this analysis has identified that indigenous people in Mexico faced worse COVID-19 outcomes than the general population and found the existence of systematic barriers that affect indigenous groups in a distinct and exclusionary manner. Hence, since COVID-19 is exacerbating the pre-existing, deep-rooted and longstanding health inequalities between indigenous and non-indigenous people, it is imperative to design programmes that prioritise and target indigenous people and to enhance the current social and health policies if the disproportionate impact of this pandemic is aimed to be mitigated.

Conclusions

This thesis concludes by summarising the key findings in relation to the research objectives and discussing the principal contributions. It also reviews the main limitations and proposes opportunities for future research. The core research objective of the thesis work was to investigate the presence, magnitude, and characteristics of health inequalities in the Mexican population. First, inequalities in nutritional outcomes were studied since Mexico faces one of the most acute and critical obesity crises globally. The argument throughout has been that obesity and overweight have often been neglected as public health problems and, to the best of our knowledge, there are no studies that have approached the obesity problem in Mexico as being the result of poor opportunities. To this end, the circumstance variables included in Chapters 1 and 2 were not chosen based on data-driven methods, nor based on the statistical criterion of exogeneity, but rather purposefully selected based on a fundamental-rights approach. Basically, focusing on the norms and laws that are supposed to guarantee an equal playing field for everyone. Thus, despite some of the variables included on the vector of circumstances being partially endogenous, they were included as they represent illegitimate and contemporaneous sources of disparities. The first paper examines the adult population at two different points in time and looks exclusively at the presence of illegitimate inequalities, thus whether inequities exist. Results pointed out the existence of inequities in a magnitude of 3-4%. However, by exploring the whole BMI and WC distribution, the beyond-the-mean analysis indicated heterogeneous levels of *ex-ante* IOp: circumstances contribute more to IOP for those people sitting in the middle of the WC and BMI distributions in 2012, and at the top of the distribution in 2018. The decomposition analysis permitted the identification of the main drivers of these inequities. In both, the mean-based and beyond-the-mean analyses, parental health conditions were found to be the main source behind the disparities. We found that the presence of diabetes and hypertension in individual's parents explained, on average, around 57%-64% of the variation in individual BMI and WC, being of higher relevance for people sitting at the *25th* percentiles of both distributions. This chapter contributes to several spheres. First, the growing literature about the identification of inequities in health approaching them via the

IOp framework. Second, the literature about the identification of inequalities at different points of the distribution, moving away from mean-based measures. This is of high relevance, given the over-nutrition outcomes studied here in which both, the bottom and top parts of the distribution are ill-related outcomes and are part of the nutrition continuum. Third, these results inform the policy-making process, by identifying that illegitimate inequalities in adult outcomes are driven by factors such as parent's health conditions suggests the need for family-based nutrition interventions, for example, to revert the long-lasting effect that exposure to obesogenic-prone behaviours during people's childhood have later in their adulthood.

Precisely, while the first chapter focused on two cross-sectional surveys and the adult population, the second chapter went further and analysed inequalities in nutritional outcomes across the life cycle, incorporating dynamic factors in the study of IOp in nutrition-related outcomes in Mexico. The late 1980s and early 1990s were difficult years for the Mexican economy. Financial debt escalated, inflation rose, purchasing power plummeted and market liberalisation commenced. This was reflected not only in changes in eating patterns but also in the population's epidemiological profile: the double burden of malnutrition originated. To account for these events, we study IOp in nutritional outcomes in a broad sense, covering under and over-nutrition as well as the whole distribution. Notwithstanding, longitudinal data covering this period do not exist for Mexico. In response, a new approach was proposed. Using matching and re-weighting techniques, we constructed a pseudo birth-cohort panel for people born between 1983 and 1988. With this data structure, covering individuals from birth until their middle thirties, we measured both *ex-ante* and *ex-post* IOp. The incorporation of the *ex-post* approach contributes to a better understanding of the mediating role of efforts in the relationship between circumstances and nutritional outcomes. The life-cycle approach also provided the opportunity to analyse how parents' efforts, along with household circumstances condition children's opportunities later in life. Results from this chapter allow us to challenge prevailing conceptions about: i) the study of under and over nutritional inequalities independently; ii) the influence of familial conditions on an individual's initial stock of health, iii) the role of morally acceptable sources of inequalities, such as peoples' efforts, on nutritional inequalities and, iv) differences in inequality magnitudes between women and men. First, we unveiled that a double burden of malnutrition exists in the Mexican adult population. For example, by 2018, 41% of the sampled individuals born between 1983 and 1988 had excess weight or adiposity and anaemia simultaneously and those circumstances explained approximately 78% of the variation in this outcome, while efforts only 22%. Second, we found that inequalities related to circumstances in under nutritional outcomes exist from childhood to adulthood in relatively similar magnitudes, despite the presence of several public programmes that targeted children's nutritional outcomes in Mexico. This seems paradoxical and further work needs to be done to investigate the long-term effects of these programmes. This chapter also challenged the conventional idea that over-nutrition, expressed commonly as overweight or obesity, is mainly driven by people's choices. Therefore, policies aimed to reduce inequalities such be compensatory rather than reward-based. Finally, this analysis found that inequalities exclusively related to circumstances were higher in girls younger than five years. However, as individuals aged, inequalities were higher in men, compared to women. When disentangling the effect of circumstances and efforts, it was found that efforts are of higher relevance in explaining inequalities in over-nutrition outcomes for women than men. These results call for a re-definition of the causes of obesity as well as re-thinking the predominant role that people's efforts, choices and behaviours have had when designing policies and programmes to tackle obesity and malnutrition.

The third chapter, although not focusing on inequalities in nutritional outcomes, showed that health inequalities in Mexico, regardless of the outcome being studied, are due to unequal access to fundamental rights. The chapter investigated the factors behind differences in COVID-19 outcomes between the general population and indigenous populations. The focus was to reflect a worsening condition, thus reflecting negative outcomes. It was found that indigenous people responded worse to the pandemic. Based on Oaxaca decomposition analyses, and accounting for individuals, as well as structural factors reflecting the relevance of the social and economic environment where people live, we explained why indigenous individuals were more affected by COVID-19, finding out that although people's underlying health conditions were the main driver of the explained differences in hospitalisations and deaths due to COVID-19, factors such as where people received care and household conditions accounted for a large part of the ethnic gap in these outcomes. This leads to several conclusions. First, indigenous people were in a worse position, in terms of health status, than the general population. This, along with other socioeconomic factors, situated them in a vulnerable position. Second, the results made explicit negative features of the Mexican health system and demonstrated that people's affiliation matter. Overall, health insurance in Mexico is not a subject of choice but mainly depends on people's jobs. Therefore, homogeneous quality of care across public health institutions is paramount to avoid illegitimate health inequalities. Third, although it is not possible to claim that the unexplained component in the Oaxaca decomposition represents discrimination, we still found discriminatory effects against indigenous people. There are no legitimate reasons to observe differences in water or electricity availability in people's households or information and communication technology accessibility since these public services have been declared as fundamental rights and are an entitlement for all. All in all, we observe a dissociation between acquired high-level commitments by the Mexican State to protect, prioritise and guarantee the welfare of indigenous communities -that have for decades been unfairly treated- and ongoing social, health and economic policies.

These essays are not without limitations, the most relevant has already been mentioned. One worth of further discussion is the lack of causal interpretation of the results. This could be a limitation if the research questions posed were to evaluate the causal effect of a policy, intervention, or programme on reducing health inequalities. But this was not an explicit part of the research objectives. One source of concern in the first chapter is that when estimating IOp, efforts, that are unobserved, could influence people's circumstances. Nevertheless, the study of early life circumstances reassures and provides more confidence against this occurring. It is also acknowledged that unobserved circumstances might also be a matter of concern. However, we are clear, in the first and second chapters, that our measurements reflect lower bounds of IOp. Alternatively, future studies that are not concerned with a right-based approach, might focus on data-driven methods to select circumstances or types. For example, recent studies have made use of machine learning algorithms, such as regression trees and forests to select circumstances (Brunori et al., 2018), latent class models (Carrieri et al., 2020) or cluster analysis (Aizawa, 2021) to identify types and model-based recursive partitioning algorithms to estimate IOp (Brunori et al., 2022).

This thesis has also opened new research questions to explore. Although the use of RIF models explores inequalities across different percentiles, future studies could measure disparities using other dissimilarity methods. Moving away from the typical mean-based perspective and putting more attention to the bottom and top parts of the distribution the nutrition continuum could be better understood. The use of polarisation indices, combined with life-course analysis could gauge nutrition-related health inequities across time. This would shed further light on the dynamics of malnutrition. Findings from this chapter also indicate that parents' health conditions are the factors that contributed the most to inequities, however, the role of epigenetics is still missing. To date, it is not clear that bad genetic luck inherited from parents to children exclusively determines health outcomes later in life nor that this unfortunate inheritance is not revertible via individual's own choices, behaviours, or efforts. Thus, data about familial genetics could help to better understand the role of luck in health.

Findings from the second chapter have also unfolded new research questions. The construction of the pseudo birth cohort of people born between 1983 to 1988 made explicit the high prevalence of malnutrition at young adulthood. This is puzzling since Mexico implemented one of the most comprehensive social programmes from 1988 to 2019, which coincides with the period studied in this analysis. This programme, which took several names: Solidaridad, Progresa, Oportunidades and Prospera, aimed to reduce poverty via cash transfers conditional on school attendance, primary health visits, and provision of nutritional food supplements and nutrition literacy. Therefore, future research could focus on exploring the long-term effect of this and other public programmes on malnutrition. This study also poses new areas of research, about effort-trajectories. Most of the research has focused on studying the evolution of circumstances on people's health outcomes. However, less attention has received effort trajectories during different stages of adulthood. One of the main critiques behind measuring IOp in health using cross-sectional data is that it penalises contemporaneous levels of efforts (Williams et al., 2000). As people change their choices, behaviours and efforts, the study of effort trajectories is relevant to better understanding inequality dynamics across the life cycle, and therefore tailoring with more precision compensation or reward policies. Finally, the proposal of using repeated cross-sections to create a pseudo panel opens the opportunity to potentially answer life-course research questions that had previously been left unanswered. For example, a 2019 study that analysed inequalities in the double burden of malnutrition among youth in Ethiopia, India, Peru, and Vietnam highlighted that the analysis was possible due to a unique longitudinal dataset that collected information over a span of 14 years (Schott et al., 2019). The lack of longitudinal data has hindered the production of policy-relevant robust evidence. Therefore, the proposed technique in this thesis could overcome this limitation.

Lastly, the study of COVID-19 inequalities contributes to the growing body of evidence regarding understanding which factors that contribute the most to the ethnic health gap. Nonetheless, the specific pathways through which socioeconomic municipal characteristics and underlying conditions lead to higher rates of COVID-19 hospitalisations and deaths among indigenous people require further research. Findings from this study suggest that the disparities among indigenous and non-indigenous people could potentially exist in other areas related to COVID-19, such as the distribution and vaccination rates. The hypothesis that epidemics in Mexico deepen inequalities between the general population and indigenous people could be further explored by taking the case of a previous flu epidemic and comparing hospitalisation and dead rates due to influenza A(H1N1), an epidemic that hit Mexico in 2009.

All in all, neglecting health inequities could translate into a serious financial burden for individuals, households, the economy, and the health system. These essays support the supposition that inequalities are predetermined and are likely to reproduce across generations should no public interventions take place. The ethical principle of compensation cannot be realised without the commitment to guarantee access to fundamental rights made for everyone, regardless of people's place of birth and living, sex or ethnicity.

References

Aitsi-Selmi, A. (2015). “Households with a Stunted Child and Obese Mother: Trends and Child Feeding Practices in a Middle-Income Country, 1992–2008”. In: *Maternal and Child Health Journal* 19.6, pp. 1284–1291. ISSN: 1092-7875. DOI: 10.1007/s10995-014-1634-5.

Aizawa, T. (Dec. 2019). “Ex-ante Inequality of Opportunity in Child Malnutrition: New Evidence from Ten Developing Countries in Asia”. In: *Economics and Human Biology* 35, pp. 144–161. ISSN: 1873-6130. DOI: 10.1016/j.ehb.2019.06.003.

— (Feb. 11, 2020). “Trajectory of inequality of opportunity in child height growth: Evidence from the Young Lives study”. In: *Demographic Research* 42.7, pp. 165–202. ISSN: 1435-9871. DOI: 10.4054/DemRes.2020.42.7.

— (Dec. 1, 2021). “Inequality of opportunity in infant mortality in South Asia: A decomposition analysis of survival data”. In: *Economics & Human Biology* 43, p. 101058. ISSN: 1570-677X. DOI: 10.1016/j.ehb.2021.101058.

Alberti, K. G. M. M. et al. (Oct. 2009). “Harmonizing the metabolic syndrome: a joint interim statement of the International Diabetes Federation Task Force on Epidemiology and Prevention; National Heart, Lung, and Blood Institute; American Heart Association; World Heart Federation; International Atherosclerosis Society; and International Association for the Study of Obesity”. eng. In: *Circulation* 120.16, pp. 1640–1645. ISSN: 1524-4539. DOI: 10.1161/CIRCULATIONAHA.109.192644.

Alcalde-Rabanal, J. E. et al. (2018). “The complex scenario of obesity, diabetes and hypertension in the area of influence of primary healthcare facilities in Mexico”. In: *PLOS ONE* 13.1, e0187028. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0187028.

Altamirano, I. et al. (2018). *Inequalities in Mexico 2018*. Mexico City: El Colegio de Mexico, Red de Estudios sobre Desigualdades, p. 144.

Apouey, B. (2007). “Measuring health polarization with self-assessed health data”. In: *Health Economics* 16.9, pp. 875–894. ISSN: 1099-1050. DOI: 10.1002/hec.1284.

Arneson, R. J. (1989). "Equality and Equal Opportunity for Welfare". In: *Philosophical Studies* 56.1, pp. 77–93. DOI: 10.1007/BF00646210.

Arroyo, P. et al. (2004). "Changes in the Household Calorie Supply during the 1994 Economic Crisis in Mexico and Its Implications on the Obesity Epidemic". In: *Nutrition Reviews* 62 (s2), S163–S168. ISSN: 1753-4887. DOI: 10.1111/j.1753-4887.2004.tb00088.x.

Assaad, R. et al. (2012). "Inequality of Opportunity in Child Health in the Arab World and Turkey". In: *Middle East Development Journal* 4.2. ISSN: 1793-8120.

Barquera, S. et al. (Dec. 2008). "Energy Intake from Beverages Is Increasing among Mexican Adolescents and Adults". en. In: *The Journal of Nutrition* 138.12, pp. 2454–2461. ISSN: 0022-3166. DOI: 10.3945/jn.108.092163.

Barquera, S. et al. (2013). "Prevalencia de obesidad en adultos mexicanos, ENSANUT 2012". In: *Salud Pública de México* 55.2, S151–S160. ISSN: 0036-3634, 1606-7916.

Barquera, S. et al. (Sept. 1, 2020). "Obesity in Mexico: rapid epidemiological transition and food industry interference in health policies". In: *The Lancet Diabetes & Endocrinology* 8.9, pp. 746–747. ISSN: 2213-8587. DOI: 10.1016/S2213-8587(20)30269-2.

Barrientos-Gutierrez, T. et al. (2017). "Expected population weight and diabetes impact of the 1-peso-per-litre tax to sugar sweetened beverages in Mexico". eng. In: *PloS One* 12.5, e0176336. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0176336.

Basto-Abreu, A. et al. (2018). "The Relationship of Socioeconomic Status with Body Mass Index Depends on the Socioeconomic Measure Used". In: *Obesity* 26.1, pp. 176–184. ISSN: 1930-739X. DOI: 10.1002/oby.22042.

Batal, M. et al. (Nov. 1, 2018). "The nutrition transition and the double burden of malnutrition". In: *Medecine Et Sante Tropicales* 28.4, pp. 345–350. ISSN: 2261-2211. DOI: 10.1684/mst.2018.0831.

Batis, C. et al. (2018). "Diet in Mexico and health effects". Spanish. In: *Obesity in Mexico State of the public policy and recommendations for its prevention and control*. First. Cuernavaca, Mexico: National Institute of Public Health.

Bell, D. (1972). "Meritocracy and Equality". English. In: *The Public Interest; New York* 29. ISSN: 0033-3557.

Beltrán-Sánchez, H. et al. (Oct. 2011). "Links between childhood and adult social circumstances and obesity and hypertension in the Mexican population". In: *Journal of Aging and Health* 23.7, pp. 1141–1165. ISSN: 1552-6887. DOI: 10.1177/0898264311422255.

Bennia, F. et al. (2022). "Is body weight better distributed among men than among women? A robust normative analysis for France, the UK, and the US*". In: *The Scandinavian Journal of Economics* 124.1, pp. 69–103. ISSN: 1467-9442. DOI: 10.1111/sjoe.12443.

Blackwell, M. et al. (Dec. 2009). "Cem: Coarsened Exact Matching in Stata". In: *The Stata Journal: Promoting communications on statistics and Stata* 9.4, pp. 524–546. ISSN: 1536-867X, 1536-8734. DOI: 10.1177/1536867X0900900402.

Blinder, A. S. (1973). "Wage Discrimination: Reduced Form and Structural Estimates". In: *The Journal of Human Resources* 8.4, pp. 436–455. ISSN: 0022-166X. DOI: 10.2307/144855.

Borgen, N. T. (2016). "Fixed effects in unconditional quantile regression". In: *Stata Journal* 16.2, pp. 403–415. DOI: 10.1177/1536867X1601600208.

Brisbois, T. D. et al. (Apr. 2012). "Early markers of adult obesity: a review". In: *Obesity Reviews* 13.4, pp. 347–367. ISSN: 1467-7881. DOI: 10.1111/j.1467-789X.2011.00965.x.

Brunori, P. (2017). "The Perception of Inequality of Opportunity in Europe". In: *Review of Income and Wealth* 63.3, pp. 464–491. ISSN: 1475-4991. DOI: 10.1111/roiw.12259.

Brunori, P. et al. (2018). *The Roots of Inequality: Estimating Inequality of Opportunity from Regression Trees.* ifo Working Paper Series 252. ifo Institute - Leibniz Institute for Economic Research at the University of Munich.

Brunori, P. et al. (Jan. 2022). *Model-based recursive partitioning to estimate unfair health inequalities in the United Kingdom Household Longitudinal Study.* LSE Research Online Documents on Economics 113538. London School of Economics and Political Science, LSE Library.

Bruoni, P. (2016). "How to measure inequality of opportunity: a hands-on guide". In: *Institute for Social Science Research, The University of Queensland.* 4th ser.

Carrieri, V. et al. (2020). "A latent class approach to inequity in health using biomarker data". In: *Health Economics* n/a (). ISSN: 1099-1050. DOI: 10.1002/hec.4022.

Cavaliere, A. et al. (2013). "Time Preference and Health: The Problem of Obesity". In: *Proceedings in Food System Dynamics* 0.0, pp. 367–380. ISSN: 2194-511X. DOI: 10.18461/pfsd.2013.1323.

Cecchini, M. et al. (Nov. 20, 2010). "Tackling of unhealthy diets, physical inactivity, and obesity: health effects and cost-effectiveness". In: *Lancet (London, England)* 376.9754, pp. 1775–1784. ISSN: 1474-547X. DOI: 10.1016/S0140-6736(10)61514-0.

Chávez -Juárez, F. W. et al. (2014). "iop: Estimating ex-ante inequality of opportunity". In: *The Stata Journal* 14.4, pp. 830–846. DOI: 10.1177/1536867X1401400408.

Clark, S. E. et al. (2012). "Exporting obesity: US farm and trade policy and the transformation of the Mexican consumer food environment". In: *International Journal of Occupational and Environmental Health* 18.1, pp. 53–64. DOI: 10.1179/1077352512Z.0000000007.

Clément, M. et al. (Mar. 1, 2021). "Does inequality have a silver lining? Municipal income inequality and obesity in Mexico". In: *Social Science & Medicine* 272, p. 113710. ISSN: 0277-9536. DOI: 10.1016/j.socscimed.2021.113710.

Comisión Nacional para el Desarrollo de los Pueblos Indigenas, C. (2015). *Indicadores Socioeconómicos de los Pueblos Indígenas de México, 2015*. Mexico City: Comisión Nacional para el Desarrollo de los pueblos Indigenas.

Consejo Nacional de Evaluación de la Política Social, C. (2007). *Anexo Metodológico del Índice de Rezago Social*.

— (2010). *Informe de evolución histórica de la situación nutricional de la población y los programas de alimentación, nutrición y abasto en México*. ISBN: 978-607-95482-5-4.

Contoyannis, P. et al. (2007). "Using relative distributions to investigate the body mass index in England and Canada". In: *Health Economics* 16.9, pp. 929–944. ISSN: 1099-1050. DOI: 10.1002/hec.1240.

Cotton, J. (1988). "On the Decomposition of Wage Differentials". In: *The Review of Economics and Statistics* 70.2, pp. 236–243. ISSN: 0034-6535. DOI: 10.2307/1928307.

Crossman, A. et al. (Nov. 2006). "The family environment and American adolescents' risk of obesity as young adults". en. In: *Social Science & Medicine* 63.9, pp. 2255–2267. ISSN: 0277-9536. DOI: 10.1016/j.socscimed.2006.05.027.

Cunha, F. et al. (May 2007). "The Technology of Skill Formation". In: *American Economic Review* 97.2, pp. 31–47. ISSN: 0002-8282. DOI: 10.1257/aer.97.2.31.

Cunha, F. et al. (2008). "Formulating, Identifying and Estimating the Technology of Cognitive and Noncognitive Skill Formation". In: *The Journal of Human Resources* 43.4. Publisher: [University of Wisconsin Press, Board of Regents of the University of Wisconsin System], pp. 738–782. ISSN: 0022-166X.

— (2009). "The Economics and Psychology of Inequality and Human Development". In: *Journal of the European Economic Association* 7.2, pp. 320–364. ISSN: 1542-4766.

Dalton, M. et al. (2003). "Waist circumference, waist–hip ratio and body mass index and their correlation with cardiovascular disease risk factors in Australian adults". In: *Journal of Internal Medicine* 254.6, pp. 555–563. ISSN: 1365-2796. DOI: 10.1111/j.1365-2796.2003.01229.x.

Dávila-Torres, J. et al. (Apr. 7, 2015). "Panorama de la obesidad en México". In: *Revista Médica del Instituto Mexicano del Seguro Social* 53.2, pp. 240–249.

Davillas, A. et al. (Jan. 1, 2020a). "Ex ante inequality of opportunity in health, decomposition and distributional analysis of biomarkers". In: *Journal of Health Economics* 69, p. 102251. ISSN: 0167-6296. DOI: 10.1016/j.jhealeco.2019.102251.

— (Aug. 2020b). "Regional inequalities in adiposity in England: distributional analysis of the contribution of individual-level characteristics and the small area obesogenic environment". en.

In: *Economics & Human Biology* 38, p. 100887. ISSN: 1570-677X. DOI: 10.1016/j.ehb.2020.100887.

Davis, M. M. (May 12, 2004). "Race-Based Immunization Recommendations and the Potential to Reduce Health Disparities". In: *JAMA* 291.18, pp. 2253–2255. ISSN: 0098-7484. DOI: 10.1001/jama.291.18.2253.

Denova-Gutiérrez, E. et al. (Dec. 2016). "Relative validity of a food frequency questionnaire to identify dietary patterns in an adult Mexican population". In: *Salud Pública de México* 58, pp. 608–616. ISSN: 0036-3634, 0036-3634. DOI: 10.21149/spm.v58i6.7842.

Deutsch, J. et al. (July 1, 2018). "Using the Shapley Decomposition to Disentangle the Impact of Circumstances and Efforts on Health Inequality". In: *Social Indicators Research* 138.2, pp. 523–543. ISSN: 1573-0921. DOI: 10.1007/s11205-017-1690-5.

Díaz de León-Martínez, L. et al. (Sept. 1, 2020). "Critical review of social, environmental and health risk factors in the Mexican indigenous population and their capacity to respond to the COVID-19". In: *Science of The Total Environment* 733, p. 139357. ISSN: 0048-9697. DOI: 10.1016/j.scitotenv.2020.139357.

Ding, L. et al. (2021). "Ex ante Inequality of Opportunity in Health among the Elderly in China: A Distributional Decomposition Analysis of Biomarkers". In: *Review of Income and Wealth* n/a. ISSN: 1475-4991. DOI: 10.1111/roiw.12514.

Doak, C. M. et al. (Jan. 2005). "The dual burden household and the nutrition transition paradox". In: *International Journal of Obesity (2005)* 29.1, pp. 129–136. ISSN: 0307-0565. DOI: 10.1038/sj.ijo.0802824.

Dobbs, R. et al. (2014). *How the world could better fight obesity*. McKinsey & Company.

Dworkin, R. (1981a). "What is Equality? Part 1: Equality of Welfare". In: *Philosophy & Public Affairs* 10.3, pp. 185–246. ISSN: 0048-3915.

— (1981b). "What is Equality? Part 2: Equality of Resources". In: *Philosophy & Public Affairs* 10.4, pp. 283–345. ISSN: 0048-3915.

ECLAC (2017). *The cost of the double burden of malnutrition: Social and economic impact - Summary of the pilot study in Chile, Ecuador and Mexico - World*. 1.

Ersado, L. et al. (Sept. 2014). *Inequality of Opportunity Among Egyptian Children*. Washington, DC: World Bank. DOI: 10.1596/1813-9450-7026.

Esposito, L. et al. (Apr. 1, 2020). "The economic gradient of obesity in Mexico: Independent predictive roles of absolute and relative wealth by gender". In: *Social Science & Medicine* 250, p. 112870. ISSN: 0277-9536. DOI: 10.1016/j.socscimed.2020.112870.

Fairlie, R. W. (1999). “The Absence of the African-American Owned Business: An Analysis of the Dynamics of Self-Employment”. In: *Journal of Labor Economics* 17.1, pp. 80–108. ISSN: 0734-306X. DOI: 10.1086/209914.

— (2005). “An extension of the Blinder-Oaxaca decomposition technique to logit and probit models”. In: *Journal of Economic and Social Measurement* 30.4, pp. 305–316. ISSN: 1875-8932. DOI: 10.3233/JEM-2005-0259.

Fajardo-Gonzalez, J. (Dec. 1, 2016). “Inequality of opportunity in adult health in Colombia”. In: *The Journal of Economic Inequality* 14.4, pp. 395–416. ISSN: 1573-8701. DOI: 10.1007/s10888-016-9338-2.

Fernald, L. C. et al. (2007). “Overweight with concurrent stunting in very young children from rural Mexico: prevalence and associated factors”. In: *European Journal of Clinical Nutrition; London* 61.5, pp. 623–32. ISSN: 09543007. DOI: http://dx.doi.org/10.1038/sj.ejcn.1602558.

Ferreira, F. H. G. et al. (2011). “The Measurement of Inequality of Opportunity: Theory and an Application to Latin America”. In: *Review of Income and Wealth* 57.4, pp. 622–657. ISSN: 1475-4991. DOI: 10.1111/j.1475-4991.2011.00467.x.

Finkelstein, E. A. et al. (May 1, 2010). “The economics of obesity”. In: *The American Journal of Clinical Nutrition* 91.5, 1520S–1524S. ISSN: 0002-9165. DOI: 10.3945/ajcn.2010.28701E.

Firpo, S. et al. (2009). “Unconditional Quantile Regressions”. en. In: *Econometrica* 77.3, pp. 953–973. ISSN: 1468-0262. DOI: 10.3982/ECTA6822.

Fleurbaey, M. et al. (2013). “Ex Ante Versus Ex Post Equality of Opportunity”. In: *Economica* 80.317.

Flint, A. J. et al. (2010). “Body mass index, waist circumference, and risk of coronary heart disease: a prospective study among men and women”. In: *Obesity research & clinical practice* 4.3, e171–e181. ISSN: 1871-403X. DOI: 10.1016/j.orcp.2010.01.001.

Fortin, N. et al. (Jan. 1, 2011). “Chapter 1 - Decomposition Methods in Economics”. In: *Handbook of Labor Economics*. Ed. by O. Ashenfelter et al. Vol. 4. Elsevier, pp. 1–102. DOI: 10.1016/S0169-7218(11)00407-2.

Frenk, J. et al. (2015). “Ethical and Human Rights Foundations of Health Policy: Lessons from Comprehensive Reform in Mexico”. In: *Health and Human Rights Journal* 17.2.

García-Peña, C. et al. (Apr. 8, 2022). “Variability in case fatality rate risk due to Covid-19 according to health services provider in Mexico City hospitals”. In: *Salud Pública de México* 64.2. Number: 2, pp. 119–130. ISSN: 1606-7916. DOI: 10.21149/12995.

Giannouchos, T. V. et al. (Jan. 1, 2020). “Characteristics and risk factors for COVID-19 diagnosis and adverse outcomes in Mexico: an analysis of 89,756 laboratory–confirmed COVID-19 cases”.

In: *European Respiratory Journal*. ISSN: 0903-1936, 1399-3003. DOI: 10.1183/13993003.02144-2020.

Giuntella, O. et al. (Jan. 1, 2020). "Weight gains from trade in foods: Evidence from Mexico". In: *Journal of International Economics* 122, p. 103277. ISSN: 0022-1996. DOI: 10.1016/j.jinteco.2019.103277.

González, A. M. et al. (Mar. 1, 2013). "Asociación entre la ingesta de calcio dietético y el índice de masa corporal elevado en adultos mexicanos de 20 a 59 años de edad: estudio de corte transversal". In: *Medwave* 13.2. ISSN: 0717-6384. DOI: 10.5867/medwave.2013.02.5635.

Grossman, M. (Mar. 1, 1972). "On the Concept of Health Capital and the Demand for Health". In: *Journal of Political Economy* 80.2, pp. 223–255. ISSN: 0022-3808. DOI: 10.1086/259880.

Gustavo, O. et al. (2006). *Encuesta Nacional de Salud y nutrición 2006*. Cuernavaca, Mexico: Instituto Nacional de Salud Pública. ISBN: 970-9874-17-9.

Gutierrez, J. P. et al. (Oct. 8, 2020). "Non-communicable diseases and inequalities increase risk of death among COVID-19 patients in Mexico". In: *PLOS ONE* 15.10, e0240394. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0240394.

Haire-Joshu, D. et al. (2016). "Preventing Obesity Across Generations: Evidence for Early Life Intervention". In: *Annual Review of Public Health* 37.1, pp. 253–271. DOI: 10.1146/annurev-publhealth-032315-021859.

Hawkins, R. B. et al. (2020). "Socio-economic status and COVID-19–related cases and fatalities". In: *Public Health* 189, pp. 129–134. ISSN: 0033-3506. DOI: https://doi.org/10.1016/j.puhe.2020.09.016.

Health, M. of (2010). *Acuerdo Nacional para la Salud Alimentaria*.

Heckman, J. J. (2012). "The developmental origins of health". In: *Health Economics* 21.1, pp. 24–29. ISSN: 1099-1050. DOI: 10.1002/hec.1802.

Hernández-Galdamez, D. R. et al. (Oct. 2020). "Increased Risk of Hospitalization and Death in Patients with COVID-19 and Pre-existing Noncommunicable Diseases and Modifiable Risk Factors in Mexico". In: *Archives of Medical Research* 51.7, pp. 683–689. ISSN: 0188-4409. DOI: 10.1016/j.arcmed.2020.07.003.

Hernández-Garduño, E. (July 1, 2020). "Obesity is the comorbidity more strongly associated for Covid-19 in Mexico. A case-control study". In: *Obesity Research & Clinical Practice* 14.4, pp. 375–379. ISSN: 1871-403X. DOI: 10.1016/j.orcp.2020.06.001.

Hojjat, T. A. et al. (2017). *The Economics of Obesity: Poverty, Income Inequality and Health*. en. SpringerBriefs in Public Health. Springer Singapore. ISBN: 978-981-10-2911-0. DOI: 10.1007/978-981-10-2911-0.

Horton, R. (Sept. 26, 2020). "Offline: COVID-19 is not a pandemic". In: *The Lancet* 396.10255, p. 874. ISSN: 0140-6736. DOI: 10.1016/S0140-6736(20)32000-6.

Ibarra-Nava, I. et al. (Mar. 10, 2021). "Ethnic disparities in COVID-19 mortality in Mexico: A cross-sectional study based on national data". In: *PLOS ONE* 16.3, e0239168. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0239168.

INEGI (2015). *Porcentaje de población que cuenta con servicio de agua entubada en su hogar 2010 y 2015*. URL: https://datos.gob.mx/busca/dataset/porcentaje-de-poblacion-que-cuenta-con-servicio-de-agua-entubada-en-su-hogar-derecho-al-medio-a (visited on 04/06/2020).

— (2020). "Presentación de resultados. Estados Unidos Mexicanos". In: p. 116.

Trabajo, O. I. del et al., eds. (2009). *Convenio Núm. 169 de la OIT sobre pueblos indígenas y tribales en países independientes: Declaración de las Naciones Unidas sobre los Derechos de los Pueblos Indígenas*. Lima: OIT. 105 pp. ISBN: 978-92-2-322580-3 978-92-2-322581-0.

Jacob, C. M. et al. (2017). *The importance of a life-course approach to health: Chronic disease risk from preconception through adolescence and adulthood: White paper*. English.

Jacobs, E. J. et al. (Aug. 9, 2010). "Waist Circumference and All-Cause Mortality in a Large US Cohort". In: *Archives of Internal Medicine* 170.15, pp. 1293–1301. ISSN: 0003-9926. DOI: 10.1001/archinternmed.2010.201.

Jann, B. (Dec. 1, 2008). "The Blinder–Oaxaca Decomposition for Linear Regression Models:" in: *The Stata Journal*. DOI: 10.1177/1536867X0800800401.

— (June 2018). "Decomposition methods in the social sciences". info:eu-repo/semantics/conferenceObject. info:eu-repo/semantics/conferenceObject. University of Bamberg. DOI: 10.7892/boris.117107.

Jarolimova, J. et al. (2013). "Obesity: Its Epidemiology, Comorbidities, and Management". In: *The Primary Care Companion for CNS Disorders* 15.5. ISSN: 2155-7772. DOI: 10.4088/PCC.12f01475.

Jones, A. M. (Apr. 13, 2019). "Equity, opportunity and health". In: *Empirica* 46, pp. 413–421. ISSN: 1573-6911. DOI: 10.1007/s10663-019-09440-x.

Jones, A. M. et al. (Oct. 1, 2014). "Equalising opportunities in health through educational policy". In: *Social Choice and Welfare* 43.3, pp. 521–545. ISSN: 1432-217X. DOI: 10.1007/s00355-014-0793-z.

Juárez-Ramírez, C. et al. (Apr. 2014). "La desigualdad en salud de grupos vulnerables de México: adultos mayores, indígenas y migrantes". In: *Revista Panamericana de Salud Pública* 35, pp. 284–290. ISSN: 1020-4989, 1020-4989, 1680-5348.

Jusot, F. et al. (2013). “Circumstances and Efforts: How Important Is Their Correlation for the Measurement of Inequality of Opportunity in Health?” en. In: *Health Economics* 22.12, pp. 1470–1495. ISSN: 1099-1050. DOI: 10.1002/hec.2896.

Jusot, F. et al. (Apr. 26, 2019). “Equality of Opportunity in Health and Healthcare”. In: *Oxford Research Encyclopedia of Economics and Finance*. DOI: 10.1093/acrefore/9780190625979.013.3.

Kanter, R. et al. (July 1, 2012). “Global Gender Disparities in Obesity: A Review”. In: *Advances in Nutrition* 3.4, pp. 491–498. ISSN: 2161-8313. DOI: 10.3945/an.112.002063.

Kral, T. V. E. et al. (July 2010). “Eating behaviors of children in the context of their family environment”. en. In: *Physiology & Behavior*. Proceedings from the 2009 Meeting of the Society for the Study of Ingestive Behavior 100.5, pp. 567–573. ISSN: 0031-9384. DOI: 10.1016/j.physbeh.2010.04.031.

Kroker-Lobos, M. F. et al. (Dec. 1, 2014). “The double burden of undernutrition and excess body weight in Mexico”. In: *The American Journal of Clinical Nutrition* 100.6, 1652S–1658S. ISSN: 0002-9165. DOI: 10.3945/ajcn.114.083832.

Lakdawalla, D. et al. (May 2002). *The Growth of Obesity and Technological Change: A Theoretical and Empirical Examination*. Working Paper 8946. National Bureau of Economic Research. DOI: 10.3386/w8946.

Lambert, F. et al. (2019). “Income Inequality and Government Transfers in Mexico”. Working Paper. International Monetary Fund. Working Paper. International Monetary Fund.

Leal, G. et al. (Dec. 2002). “Tres momentos en la política de salud y seguridad social en México”. In: *Papeles de población* 8.34, pp. 107–133. ISSN: 1405-7425.

Lefranc, A. et al. (Dec. 1, 2009). “Equality of opportunity and luck: Definitions and testable conditions, with an application to income in France”. In: *Journal of Public Economics* 93.11, pp. 1189–1207. ISSN: 0047-2727. DOI: 10.1016/j.jpubeco.2009.07.008.

León-Cortés, J. L. et al. (Feb. 7, 2019). “Health reform in Mexico: governance and potential outcomes”. In: *International Journal for Equity in Health* 18, p. 30. ISSN: 1475-9276. DOI: 10.1186/s12939-019-0929-y.

Levasseur, P. (July 1, 2015). “Causal effects of socioeconomic status on central adiposity risks: Evidence using panel data from urban Mexico”. In: *Social Science & Medicine* 136-137, pp. 165–174. ISSN: 0277-9536. DOI: 10.1016/j.socscimed.2015.05.018.

Leyva-Flores, R. et al. (Aug. 2014). “Primary Health Care Utilization by the Mexican Indigenous Population: The Role of the Seguro Popular in Socially Inequitable Contexts”. In: *PLoS ONE* 9.8. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0102781.

Leyva-Flores, R. et al. (2013). “Inequidad persistente en salud y acceso a los servicios para los pueblos indígenas de México, 2006-2012”. In: *Salud Pública de México* 55, S123–S128. ISSN: 0036-3634.

Liu, X. et al. (June 1, 2022). “Inequality of opportunity in children’s nutritional outcomes in China”. In: *Global Food Security* 33, p. 100635. ISSN: 2211-9124. DOI: 10.1016/j.gfs.2022.100635.

Long-Dunlap, K. et al. (Mar. 14, 1995). “Feeding patterns of mexican infants recorded in the 1988 national nutrition survey”. In: *Salud Pública de México* 37.2, pp. 120–129. ISSN: 1606-7916.

López, O. et al. (July 14, 2018). “In town with little water, coca-cola is everywhere. So is diabetes.” In: *The New York Times* 2018.

Mancilla-Galindo, J. et al. (Nov. 26, 2020). “Development and validation of the patient history COVID-19 (PH-Covid19) scoring system: a multivariable prediction model of death in Mexican patients with COVID-19”. In: *Epidemiology and Infection* 148. ISSN: 0950-2688. DOI: 10.1017/S0950268820002903.

Martinez, H. et al. (Mar. 14, 1995). “Anemia en mujeres de edad reproductiva. Resultados de una encuesta probabilística nacional”. In: *Salud Pública de México* 37.2, pp. 108–119. ISSN: 1606-7916.

Matthews, S. et al. (Jan. 1, 1999). “Social inequalities in health: are there gender differences?” In: *Social Science & Medicine* 48.1, pp. 49–60. ISSN: 0277-9536. DOI: 10.1016/S0277-9536(98)00288-3.

McCormack, G. R. et al. (Jan. 2011). “Associations between familial affluence and obesity risk behaviours among children”. en. In: *Paediatrics & Child Health* 16.1, pp. 19–24. ISSN: 1205-7088. DOI: 10.1093/pch/16.1.19.

McIntyre, P. B. et al. (2005). “Immunisation: reducing health inequality for Indigenous Australians”. In: *Medical Journal of Australia* 182.5, pp. 207–208. ISSN: 1326-5377. DOI: 10.5694/j.1326-5377.2005.tb06667.x.

Medina, C. et al. (Nov. 2013). “Physical inactivity prevalence and trends among Mexican adults: results from the National Health and Nutrition Survey (ENSANUT) 2006 and 2012”. In: *BMC Public Health* 13, p. 1063. ISSN: 1471-2458. DOI: 10.1186/1471-2458-13-1063.

Mexican Constitution, M. (2017). *Political Constitution of the United Mexican States*. 2017th ed.

Meza, R. et al. (Dec. 2015). “Burden of Type 2 Diabetes in Mexico: Past, Current and Future Prevalence and Incidence Rates”. In: *Preventive medicine* 81, pp. 445–450. ISSN: 0091-7435. DOI: 10.1016/j.ypmed.2015.10.015.

Monroy-Gómez-Franco, L. et al. (Sept. 24, 2020). *A Land of Unequal Chances: Social Mobility Across Mexican Regions*.

Monroy-Gómez-Franco, L. et al. (Dec. 2021). "Skin Tone Differences in Social Mobility in Mexico: Are We Forgetting Regional Variance?" In: *Journal of Economics, Race, and Policy* 4.4, pp. 257–274. DOI: 10.1007/s41996-020-00062-.

Monterrubio-Flores, E. et al. (Sept. 25, 2020). "Impact of a Double Epidemic in Mexico: Non-Communicable Diseases Increase the Case Fatality Rate with Covid-19". In: *Research Square. Preprint.* DOI: 10.21203/rs.3.rs-80669/v1.

National Council for the Evaluation of Social Development Policy (2018). *Población indígena con carencias en todos sus derechos sociales.* Mexico City: National council for the evaluation of social development policy.

National Institute of Public Health, M. (2016). *ENSANUT Medio camino 2016.*

— (2018). *ENSANUT 2018. Results.*

Neumark, D. (1988). "Employers' Discriminatory Behavior and the Estimation of Wage Discrimination". In: *The Journal of Human Resources* 23.3, pp. 279–295. ISSN: 0022-166X. DOI: 10.2307/145830.

Nie, P. et al. (2020). "Inequality of Opportunity in Bodyweight among Middle-Aged and Older Chinese: A Distributional Approach". In: *IZA Institute of Labor Economics* (DP No. 13421), p. 39.

Nielsen, L. A. et al. (Oct. 2015). "The Impact of Familial Predisposition to Obesity and Cardiovascular Disease on Childhood Obesity". In: *Obesity Facts* 8.5, pp. 319–328. ISSN: 1662-4025. DOI: 10.1159/000441375.

O'Donnell, O. et al. (2007). *Analyzing Health Equity Using Household Survey Data. A guide to techniques and their implementation.* The World Bank.

Oaxaca, R. et al. (1999). "Identification in Detailed Wage Decompositions". In: *The Review of Economics and Statistics* 81.1, pp. 154–157.

Obesity Institute of Medicine, O. (1995). *Prevention of Obesity.* National Academies Press (US).

OECD (Nov. 2011). *Diabetes prevalence and incidence.* en. Tech. rep. Paris: OECD, pp. 42–43. DOI: 10.1787/health_glance-2011-13-en.

Paes de Barros, R. et al. (Nov. 12, 2008). *Measuring inequality of opportunities in Latin America and the Caribbean.* 46827. The World Bank, pp. 1–222.

PanAmerican Health Organization, W. (2015). *Ultra-processed food and drink products in Latin America: Trends, impact on obesity, policy implications.* Tech. rep. Washington, DC.

Pérez Ferrer, C. (July 28, 2015). "Socioeconomic inequalities in obesity among Mexican adults 1988-2012". Doctoral Thesis. UCL (University College London).

Plassot, T. et al. (Sept. 2, 2022). “Inequality of Opportunity in Mexico and its Regions: A Data-Driven Approach”. In: *The Journal of Development Studies* 58.9. Publisher: Routledge _eprint: https://doi.org/10.1080/00220388.2022.2055465, pp. 1857–1873. ISSN: 0022-0388. DOI: 10.1080/00220388.2022.2055465.

Pliego, J. T. P. (Aug. 15, 2019). “Dulce exterminio: refresco y cerveza como causa desencadenante y complicaciones de la diabetes en mayas de Chiapas, México / Sweet extermination: Soda and beer, as trigger cause and complications in diabetics, among high land mayans of Chiapas, Mexico.” In: *Medicina Social* 12.2, pp. 87–95. ISSN: 1557-7112.

Popkin, B. M. (Apr. 1, 2001). “The Nutrition Transition and Obesity in the Developing World”. In: *The Journal of Nutrition* 131.3, 871S–873S. ISSN: 0022-3166. DOI: 10.1093/jn/131.3.871S.

— (Sept. 2015). “Nutrition Transition and the Global Diabetes Epidemic”. In: *Current diabetes reports* 15.9, p. 64. ISSN: 1534-4827. DOI: 10.1007/s11892-015-0631-4.

Popkin, B. M. et al. (Jan. 2012). “NOW AND THEN: The Global Nutrition Transition: The Pandemic of Obesity in Developing Countries”. In: *Nutrition Reviews* 70.1, pp. 3–21. ISSN: 0029-6643. DOI: 10.1111/j.1753-4887.2011.00456.x.

Porro, G. et al. (2009). “Random Recursive Partitioning: a matching method for the estimation of the average treatment effect”. In: *Journal of Applied Econometrics* 24.1, pp. 163–185. ISSN: 1099-1255. DOI: 10.1002/jae.1026.

Powers, D. A. et al. (2000). *Statistical Methods for Categorical Data Analysis*. 2nd Edition. San Diego, California, USA: Academic Press. ISBN: 0-12-563736-5.

Powers, D. A. et al. (Dec. 1, 2011). “Mvdcmp: Multivariate Decomposition for Nonlinear Response Models”. In: *The Stata Journal* 11.4, pp. 556–576. ISSN: 1536-867X. DOI: 10.1177/1536867X1201100404.

Puig, A. et al. (2009). “Assessing Quality across Health Care Subsystems in Mexico”. In: *The Journal of Ambulatory Care Management* 32.2, pp. 123–131. ISSN: 0148-9917. DOI: 10.1097/JAC.0b013e31819942e5.

Quezada, A. D. et al. (Dec. 16, 2015). “Time trends and sex differences in associations between socioeconomic status indicators and overweight-obesity in Mexico (2006–2012)”. In: *BMC Public Health* 15. ISSN: 1471-2458. DOI: 10.1186/s12889-015-2608-2.

Rahimi, E. et al. (Aug. 2021). “A detailed explanation and graphical representation of the Blinder-Oaxaca decomposition method with its application in health inequalities”. In: *Emerging Themes in Epidemiology* 18.1, p. 12. ISSN: 1742-7622. DOI: 10.1186/s12982-021-00100-9.

Ramos, X. et al. (2016). "Approaches to Inequality of Opportunity: Principles, Measures and Evidence". In: *Journal of Economic Surveys* 30.5, pp. 855–883. ISSN: 1467-6419. DOI: 10.1111/joes.12121.

Rangel Garrocho, C. et al. (2014). "Estructura profunda de los flujos migratorios en México, 1990-2010". In: *La situación demográfica de México 2014*. Consejo Nacional de Población, p. 32.

Rashad, I. et al. (2004). "The economics of obesity". In: *National Affairs* 156.

Rawls, J. (1971). *A theory of justice*. 1st. Hardvard University Press.

Resano-Pérez, E. et al. (2003). "Methods of the National Nutrition Survey 1999". In: *Salud Publica De Mexico* 45 Suppl 4, S558–564. ISSN: 0036-3634. DOI: 10.1590/s0036-36342003001000012.

Reyes, H. et al. (Nov. 2004). "The family as a determinant of stunting in children living in conditions of extreme poverty: a case-control study". In: *BMC Public Health* 4.1, p. 57. ISSN: 1471-2458. DOI: 10.1186/1471-2458-4-57.

Rivera-Dommarco, J. A. et al. (Mar. 14, 1995). "Déficit de talla y emaciación en menores de cinco años en distintas regiones y estratos en México". In: *Salud Pública de México* 37.2, pp. 95–107. ISSN: 1606-7916.

Rivera-Dommarco, J. A. et al. (Feb. 2010). "Effectiveness of a large-scale iron-fortified milk distribution program on anemia and iron deficiency in low-income young children in Mexico". In: *The American Journal of Clinical Nutrition* 91.2, pp. 431–439. ISSN: 1938-3207. DOI: 10.3945/ajcn.2009.28104.

Rivera-Dommarco, J. A. et al. (2018). "Position. Recommendations for a policy of State for the prevention and control of obesity in Mexico in the period 2018-2024". Spanish. In: *Obesity in Mexico State of the public policy and recommendations for its prevention and control*. First. Cuernavaca, Mexico: National Institute of Public Health.

Roemer, J. E. (1998). *Equality of opportunity*. Hardvard University Press.

— (2002). "Equality of opportunity: A progress report". In: *Social Choice and Welfare* 19.2, pp. 455–471. ISSN: 0176-1714.

Roemer, J. E. et al. (Dec. 2016). "Equality of Opportunity: Theory and Measurement". en. In: *Journal of Economic Literature* 54.4, pp. 1288–1332. ISSN: 0022-0515. DOI: 10.1257/jel.20151206.

Romero-Martínez, M. et al. (2013). "Encuesta Nacional de Salud y Nutrición 2012: diseño y cobertura". In: *Salud Pública de México* 55, S332–S340. ISSN: 0036-3634.

Romero-Martínez, M. et al. (June 2017). "Diseño metodológico de la Encuesta Nacional de Salud y Nutrición de Medio Camino 2016". In: *Salud Pública de México* 59.3, pp. 299–305. ISSN: 0036-3634. DOI: 10.21149/8593.

Romero-Martínez, M. et al. (Dec. 5, 2019). "Encuesta Nacional de Salud y Nutrición (Ensanut 2018): metodología y perspectivas". In: *Salud Pública de México* 61.6, pp. 917–923. ISSN: 1606-7916. DOI: 10.21149/11095.

Rosa Dias, P. (2009). "Inequality of opportunity in health: evidence from a UK cohort study". In: *Health Economics* 18, pp. 1057–1074.

Rosenbaum, J. E. (Jan. 1, 1995). "Changing the geography of opportunity by expanding residential choice: Lessons from the Gautreaux program". In: *Housing Policy Debate* 6.1, pp. 231–269. ISSN: 1051-1482. DOI: 10.1080/10511482.1995.9521186.

Rtveladze, K. et al. (Jan. 2014). "Obesity prevalence in Mexico: impact on health and economic burden". In: *Public Health Nutrition* 17.1, pp. 233–239. ISSN: 1475-2727. DOI: 10.1017/S1368980013000086.

Sachs, J. D. et al. (Sept. 14, 2020). "Lancet COVID-19 Commission Statement on the occasion of the 75th session of the UN General Assembly". In: *The Lancet* 0.0. Publisher: Elsevier. ISSN: 0140-6736, 1474-547X. DOI: 10.1016/S0140-6736(20)31927-9.

Saeedi, P. et al. (Nov. 1, 2019). "Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: Results from the International Diabetes Federation Diabetes Atlas, 9th edition". In: *Diabetes Research and Clinical Practice* 157. Publisher: Elsevier. ISSN: 0168-8227, 1872-8227. DOI: 10.1016/j.diabres.2019.107843.

Sánchez T., M. (2020). *La letalidad hospitalaria por covid-19 en México: desigualdades institucionales*. Taller de datos. URL: https://datos.nexos.com.mx (visited on 02/10/2021).

Sanoussi, Y. et al. (Aug. 25, 2020). "Assessing and decomposing inequality of opportunity in access to child health and nutrition in sub-Saharan Africa: evidence from three countries with low human development index". In: *International Journal for Equity in Health* 19.1, p. 143. ISSN: 1475-9276. DOI: 10.1186/s12939-020-01258-5.

Schott, W. et al. (Aug. 1, 2019). "The double burden of malnutrition among youth: Trajectories and inequalities in four emerging economies". In: *Economics & Human Biology*. New Findings in Economics and Human Biology:A special Issue in Honor of Founding Editor John Komlos 34, pp. 80–91. ISSN: 1570-677X. DOI: 10.1016/j.ehb.2019.05.009.

Servan-Mori, E. et al. (Feb. 2014). "An explanatory analysis of economic and health inequality changes among Mexican indigenous people, 2000-2010". In: *International Journal for Equity in Health* 13, p. 21. ISSN: 1475-9276. DOI: 10.1186/1475-9276-13-21.

Serván-Mori, E. et al. (July 14, 2021). "Hospitalisation and mortality from COVID-19 in Mexican indigenous people: a cross-sectional observational study". In: *J Epidemiol Community Health*. ISSN: 0143-005X, 1470-2738. DOI: 10.1136/jech-2020-216129.

Shrimpton, R. et al. (2012). *The Double Burden of Malnutrition : A Review of Global Evidence*. World Bank, Washington, DC. World Bank.

Silveira, K. B. R. et al. (June 2010). "Associação entre desnutrição em crianças moradoras de favelas, estado nutricional materno e fatores socioambientais". pt. In: *Jornal de Pediatria* 86.3, pp. 215–220. ISSN: 0021-7557. DOI: 10.1590/S0021-75572010000300009.

Słoczyński, T. (2015). "The Oaxaca–Blinder Unexplained Component as a Treatment Effects Estimator". In: *Oxford Bulletin of Economics and Statistics* 77.4, pp. 588–604. ISSN: 1468-0084. DOI: 10.1111/obes.12075.

— (May 1, 2020). "Average Gaps and Oaxaca–Blinder Decompositions: A Cautionary Tale about Regression Estimates of Racial Differences in Labor Market Outcomes". In: *ILR Review* 73.3, pp. 705–729. ISSN: 0019-7939. DOI: 10.1177/0019793919874063.

Soares, G. H. et al. (Sept. 28, 2021). "Disparities in Excess Mortality Between Indigenous and Non-Indigenous Brazilians in 2020: Measuring the Effects of the COVID-19 Pandemic". In: *Journal of Racial and Ethnic Health Disparities*. ISSN: 2196-8837. DOI: 10.1007/s40615-021-01162-w.

Soberanes, R. (Jan. 21, 2021). *Las muertes ocultas de los indígenas chiapanecos que temen al Estado*. Mexico: Mexicanos contra la corrupción y la impunidad.

Sobrino, J. (2010). *Migración interna en México durante el siglo XX*. 1. ed. México, D.F: Consejo Nacional de Población. 171 pp. ISBN: 978-970-628-961-2.

Sudharsanan, N. et al. (Apr. 1, 2019). "Impact of Coming Demographic Changes on the Number of Adults in Need of Care for Hypertension in Brazil, China, India, Indonesia, Mexico, and South Africa". In: *Hypertension* 73.4, pp. 770–776. DOI: 10.1161/HYPERTENSIONAHA.118.12337.

Sullivan, K. M. et al. (2008). "Haemoglobin adjustments to define anaemia". In: *Tropical Medicine & International Health* 13.10, pp. 1267–1271. ISSN: 1365-3156. DOI: https://doi.org/10.1111/j.1365-3156.2008.02143.x.

Tai, D. et al. (2020). "The Disproportionate Impact of COVID-19 on Racial and Ethnic Minorities in the United States". In: *Clinical Infectious Diseases* (). DOI: 10.1093/cid/ciaa815.

Téllez-Rojo, M. M. et al. (2019). "Influence of post-partum BMI change on childhood obesity and energy intake". In: *PLOS ONE* 14.12, e0224830. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0224830.

Teran-Garcia, M. et al. (2013). "FTO genotype is associated with body mass index and waist circumference in Mexican young adults". In: *Open Journal of Genetics* 03.1, pp. 44–48. ISSN: 2162-4453, 2162-4461. DOI: 10.4236/ojgen.2013.31005.

Thurstans, S. et al. (Dec. 1, 2020). "Boys are more likely to be undernourished than girls: a systematic review and meta-analysis of sex differences in undernutrition". In: *BMJ Global Health*

5.12. Publisher: BMJ Specialist Journals Section: Original research, e004030. ISSN: 2059-7908. DOI: 10.1136/bmjgh-2020-004030.

Townsend P., D. N. (1982). *Inequalities in Health: The Black Report*. Penguin, Harmondsworth.

Trannoy, A. et al. (Aug. 2010). "Inequality of opportunities in health in France: a first pass". In: *Health Economics* 19.8, pp. 921–938. ISSN: 1099-1050. DOI: 10.1002/hec.1528.

Tzioumis, E. et al. (June 2014). "Childhood dual burden of under- and over-nutrition in low- and middle-income countries: a critical review". In: *Food and nutrition bulletin* 35.2, pp. 230–243. ISSN: 0379-5721.

Ullmann, S. H. et al. (June 2011). "Socioeconomic differences in obesity among Mexican adolescents". In: *International journal of pediatric obesity : IJPO : an official journal of the International Association for the Study of Obesity* 6.2, e373–e380. ISSN: 1747-7166. DOI: 10.3109/17477166.2010.498520.

Vazquez, G. et al. (Jan. 1, 2007). "Comparison of Body Mass Index, Waist Circumference, and Waist/Hip Ratio in Predicting Incident Diabetes: A Meta-Analysis". In: *Epidemiologic Reviews* 29.1, pp. 115–128. ISSN: 0193-936X. DOI: 10.1093/epirev/mxm008.

Villalpando, S. et al. (Oct. 2006). "Fortifying milk with ferrous gluconate and zinc oxide in a public nutrition program reduced the prevalence of anemia in toddlers". In: *The Journal of Nutrition* 136.10, pp. 2633–2637. ISSN: 0022-3166. DOI: 10.1093/jn/136.10.2633.

WHO (1995). *WHO | Physical status: the use and interpretation of anthropometry*. Tech. rep.

— (2012). *Haemoglobin concentrations for the diagnosis of anaemia and assessment of severity. Vitamin and Mineral Nutrition Information System.*

— (2017a). *Double burden of malnutrition. Policy brief.* Geneva, Switzerland: World Health Organization.

— (2017b). *Primary health care systems (PRIMASYS): case study from Mexico*. Geneva, Switzerland: World Health Organization.

— (2020). *Cut-off points and summary statistics*. Global Database on Child Growth and Malnutrition. URL: https://www.who.int/nutgrowthdb/about/introduction/en/index5.html (visited on 04/02/2020).

Williams, A. et al. (Jan. 1, 2000). "Chapter 35 Equity in health". In: *Handbook of Health Economics*. Vol. 1. Elsevier, pp. 1863–1910. DOI: 10.1016/S1574-0064(00)80048-7.

Wrotniak, B. H. et al. (Apr. 2004). "Parent Weight Change as a Predictor of Child Weight Change in Family-Based Behavioral Obesity Treatment". en. In: *Archives of Pediatrics & Adolescent Medicine* 158.4, pp. 342–347. ISSN: 1072-4710. DOI: 10.1001/archpedi.158.4.342.

Yashadhana, A. et al. (Aug. 1, 2020). "Indigenous Australians at increased risk of COVID-19 due to existing health and socioeconomic inequities". In: *The Lancet Regional Health – Western Pacific* 1. ISSN: 2666-6065. DOI: 10.1016/j.lanwpc.2020.100007.

Yun, M.-S. (Feb. 1, 2004). "Decomposing differences in the first moment". In: *Economics Letters* 82.2, pp. 275–280. ISSN: 0165-1765. DOI: 10.1016/j.econlet.2003.09.008.

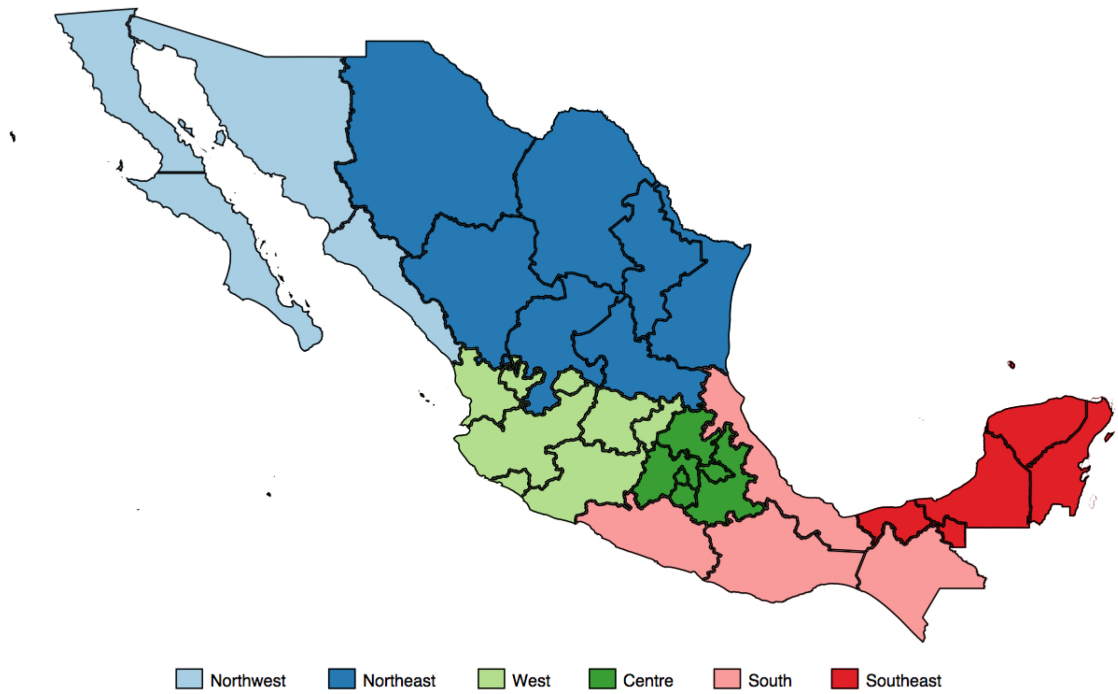
— (2005). "A Simple Solution to the Identification Problem in Detailed Wage Decompositions". In: *Economic Inquiry* 43.4, pp. 766–772. ISSN: 1465-7295. DOI: https://doi.org/10.1093/ei/cbi053.

Zhang, L. et al. (Jan. 2008). "Obesity and time preference: the health consequences of discounting the future". In: *Journal of Biosocial Science* 40.1, pp. 97–113. ISSN: 0021-9320. DOI: 10.1017/S0021932007002039.

Appendix A

Chapter 1

Figure A.0.1: Regional categorisation of the 32 Federal States of Mexico



A.1 Variable definitions

Table A.1.1: Definition of variables used as circumstances

Circumstance variables	**Type of variable**	**Definition**
Ethnicity	Binary	1 if non-indigenous, 0 if indigenous
Health insurance	Binary	1 if no health insurance, 0 otherwise 1 if public insurance (IMSS), 0 otherwise 1 if public insurance (ISSSTE), 0 otherwise 1 if public insurance (Seguro Popular), 0 otherwise 1 if public insurance (PEMEX, Secretaría de la Defensa Nacional and Secretaría de Marina), 0 otherwise 1 if private health insurance, 0 otherwise
Parental diabetes and hypertension	Binary	1 if father non-diabetic, 0 otherwise 1 if father without hypertension, 0 otherwise 1 if mother non-diabetic, 0 otherwise 1 if mother without hypertension, 0 otherwise
Running water	Binary	1 if piped water is available inside of the household, 0 otherwise 1 if piped water is available but outside of the household (e.g from public wells, nearby rivers, lakes or ponds), 0 otherwise 1 if no piped water is available either inside or outside the household, 0 otherwise
Social deprivation State-level	Binary	1 if very high State deprivation, 0 otherwise 1 if high State deprivation, 0 otherwise 1 if medium State deprivation, 0 otherwise 1 if low State deprivation, 0 otherwise 1 if very low State deprivation, 0 otherwise The categorisation results from getting quintiles of the State deprivation index, which includes the following State characteristics: percentage of the population older than 15 years old and deemed illiterate percentage of the population aged 6 to 14 who do not attend school percentage of households with individuals aged 15 to 29 that have less than 9 years of education percentage of the population older than 15 years with incomplete basic education percentage of the population without health insurance percentage of households with no floor average occupants per bedroom percentage of households without a toilet percentage of households without piped water from the public network percentage of households without sewage percentage of households without electricity percentage of households without a washing machine percentage of households without a fridge
Geographical region	Binary	1 if the person lives in the Northwest, 0 otherwise 1 if the person lives in the Northeast, 0 otherwise 1 if the person lives in the West, 0 otherwise 1 if the person lives in the Centre, 0 otherwise 1 if the person lives in the South, 0 otherwise 1 if the person lives in the Southeast, 0 otherwise Northwest: Baja California, Baja California Sur, Sinaloa and Sonora Northeast: Coahuila, Nuevo León, Tamaulipas, Chihuahua, Durango, Zacatecas and San Luis Potosí West: Aguascalientes, Colima, Guanajuato, Jalisco, Michoacán, Nayarit and Queretaro Centre: Mexico City, State of México, Hidalgo, Morelos, Puebla and Tlaxcala South: Guerrero, Oaxaca, Chiapas and Veracruz Southeast: Campeche, Quintana Roo, Tabasco and Yucatán}
Urbanity	Binary	1 if urban-metropolitan, 0 rural

A.2 Regression results

Table A.2.1: Linear regression results for all outcomes and years

	BMI 2012	**BMI 2018**	**WC 2012**	**WC 2018**
Ethnicity				
Non indigenous	0.96***	0.75***	2.83***	1.92***
	(0.15)	(0.20)	(0.35)	(0.52)
Health Affiliation				
IMSS	-0.24+	0.19	-0.24	0.77
	(0.15)	(0.24)	(0.36)	(0.61)
ISSSTE	-0.43*	-0.33	-0.41	-0.61
	(0.20)	(0.32)	(0.48)	(0.76)
Seg.Pop.	-0.11	0.40+	-0.07	1.02+
	(0.14)	(0.24)	(0.34)	(0.59)
PDM	-0.46	1.43	0.19	1.26
	(0.53)	(1.38)	(1.32)	(2.79)
Private	0.10	-0.10	-0.68	0.57
	(0.72)	(0.66)	(1.88)	(1.41)
Parents' health				
Father not Diabetic	-0.86***	-1.22***	-2.23***	-3.11***
	(0.15)	(0.21)	(0.37)	(0.49)
Father without Hypertension	-0.63***	-0.20	-1.17**	-0.34
	(0.15)	(0.20)	(0.36)	(0.47)
Mother not Diabetic.	-0.88***	-0.49**	-1.96***	-1.24**
	(0.12)	(0.18)	(0.31)	(0.43)
Mother without Hypertension	-0.34**	-0.45**	-0.70*	-0.62
	(0.11)	(0.16)	(0.28)	(0.39)
Water availability				
Piped water outside	0.06	-0.02	0.15	-0.42
	(0.12)	(0.20)	(0.31)	(0.49)
No Piped water	-0.36+	-0.54+	-1.11*	-2.07*
	(0.19)	(0.32)	(0.45)	(0.82)
State Depriv.				
High Deprivation	-0.10	0.85*	0.08	4.18***
	(0.22)	(0.38)	(0.52)	(1.27)
Medium Deprivation	-0.54*	0.76+	-0.34	4.86***
	(0.26)	(0.43)	(0.62)	(1.34)
Low Deprivation	-0.26	1.13*	0.53	5.37***
	(0.27)	(0.44)	(0.65)	(1.32)
Very Low Deprivation	0.06	1.32**	1.04	5.54***
	(0.29)	(0.49)	(0.69)	(1.45)
Geo. Region				
Urban-Metrop.	0.67***	0.60***	1.48***	1.23**
	(0.10)	(0.15)	(0.25)	(0.39)
Northeast	-0.51**	-1.07***	-0.36	-2.66***
	(0.17)	(0.29)	(0.40)	(0.80)
West	-0.68***	-1.45***	-1.06*	-3.12***
	(0.18)	(0.28)	(0.43)	(0.76)
Centre	-0.87***	-1.29***	-1.84***	-3.37***
	(0.18)	(0.31)	(0.45)	(0.83)
South	-0.66*	-0.33	-1.85**	0.84
	(0.27)	(0.47)	(0.64)	(1.51)
Southeast	1.03***	0.72*	0.42	-2.37**
	(0.21)	(0.31)	(0.50)	(0.81)
_cons	30.05***	29.41***	95.57***	94.87***
	(0.40)	(0.62)	(0.96)	(1.89)
N	27,612	12,644	26,808	12,392
r2	.111	.115	.138	.128

Notes: standard errors in parenthesis. r2=R squared
+ p<0.1, * p<0.05, ** p<0.01, *** p<0.001
Seg.Pop.=Seguro Popular. PDM=Pemex, Defensa and Marina
Very high deprivation,no health insurance, Northwest region,
piped water in household and rural area used as reference categories

Table A.2.2: Linear regression results for BMI across different percentiles for 2012

	BMI			
	q25	**q50**	**q75**	**q95**
Ethnicity				
Non indigenous	0.62***	0.95***	1.35***	1.60***
	(0.16)	(0.15)	(0.19)	(0.42)
Health Affiliation				
IMSS	0.26**	0.06	-0.48***	-1.08***
	(0.10)	(0.10)	(0.13)	(0.27)
ISSSTE	0.25	-0.37*	-0.82***	-1.69***
	(0.17)	(0.17)	(0.21)	(0.45)
Seg.Pop.	0.01	-0.09	-0.09	-0.35
	(0.10)	(0.10)	(0.12)	(0.27)
PDM	-0.80+	-0.42	0.06	-2.05+
	(0.44)	(0.43)	(0.54)	(1.18)
Private	-0.44	-0.29	-0.37	1.11
	(0.36)	(0.36)	(0.45)	(0.98)
Parents' health				
Father not Diabetic	-0.51***	-0.76***	-1.07***	-1.97***
	(0.10)	(0.10)	(0.12)	(0.27)
Father without Hypertension	-0.54***	-0.61***	-0.59***	-0.92***
	(0.10)	(0.10)	(0.12)	(0.27)
Mother not Diabetic.	-0.66***	-0.81***	-0.96***	-1.25***
	(0.09)	(0.09)	(0.11)	(0.24)
Mother without Hypertension	-0.26**	-0.28***	-0.38***	-0.93***
	(0.08)	(0.08)	(0.10)	(0.22)
Water availability				
Piped water outside	0.04	0.13	0.11	0.11
	(0.10)	(0.09)	(0.12)	(0.25)
No Piped water	-0.66***	-0.57***	-0.31	0.24
	(0.16)	(0.16)	(0.20)	(0.43)
State Depriv.				
High Deprivation	-0.03	-0.13	0.01	-0.74
	(0.18)	(0.18)	(0.22)	(0.48)
Medium Deprivation	-0.47*	-0.51*	-0.38	-1.44*
	(0.22)	(0.22)	(0.28)	(0.60)
Low Deprivation	-0.34	-0.28	-0.15	-0.70
	(0.22)	(0.21)	(0.27)	(0.58)
Very Low Deprivation	-0.06	0.02	0.47	-0.60
	(0.23)	(0.22)	(0.28)	(0.61)
Geo. Region				
Urban-Metrop.	0.51***	0.68***	0.72***	1.20***
	(0.10)	(0.10)	(0.12)	(0.27)
Northeast	-0.44**	-0.42**	-0.64***	-0.99*
	(0.15)	(0.15)	(0.19)	(0.41)
West	-0.45**	-0.59***	-0.70***	-1.78***
	(0.16)	(0.15)	(0.20)	(0.42)
Centre	-0.23	-0.60***	-1.22***	-2.25***
	(0.15)	(0.15)	(0.18)	(0.40)
South	-0.25	-0.49*	-0.95***	-2.13***
	(0.22)	(0.21)	(0.27)	(0.58)
Southeast	1.20***	1.07***	1.03***	0.51
	(0.22)	(0.21)	(0.27)	(0.58)
_cons	25.64***	29.08***	33.00***	42.64***
	(0.33)	(0.32)	(0.41)	(0.88)
N	27,612	27,612	27,612	27,612
r2	.1	.0835	.0637	.0299

Notes: standard errors in parenthesis. r2=R squared
+ p<0.1, * p<0.05, ** p<0.01, *** p<0.001
Seg.Pop.=Seguro Popular. PDM=Pemex, Defensa and Marina
Very high deprivation,no health insurance, Northwest region,
piped water in household and rural area used as reference categories

Table A.2.3: Linear regression results for WC across different percentiles for 2012

	WC			
	q25	**q50**	**q75**	**q95**
Ethnicity				
Non indigenous	2.60***	3.19***	3.69***	4.26***
	(0.44)	(0.41)	(0.48)	(0.92)
Health Affiliation				
IMSS	0.51+	-0.32	-0.58+	-1.84**
	(0.28)	(0.26)	(0.30)	(0.59)
ISSSTE	1.17*	-0.50	-1.83***	-2.63**
	(0.47)	(0.43)	(0.51)	(0.98)
Seg.Pop.	0.44	-0.06	-0.55+	-1.74**
	(0.28)	(0.26)	(0.30)	(0.58)
PDM	0.02	-0.45	-2.44+	1.90
	(1.21)	(1.12)	(1.31)	(2.53)
Private	-2.13*	-1.79+	0.65	3.74+
	(1.03)	(0.96)	(1.12)	(2.16)
Parents' health				
Father not Diabetic	-1.56***	-2.30***	-2.56***	-2.84***
	(0.28)	(0.26)	(0.30)	(0.58)
Father without Hypertension	-0.89**	-0.85***	-1.53***	-2.53***
	(0.28)	(0.26)	(0.30)	(0.57)
Mother not Diabetic.	-1.90***	-2.11***	-1.88***	-2.34***
	(0.25)	(0.23)	(0.27)	(0.51)
Mother without Hypertension	-0.68**	-0.78***	-1.10***	-1.30**
	(0.23)	(0.21)	(0.25)	(0.47)
Water availability				
Piped water outside	0.52*	-0.07	0.12	0.23
	(0.27)	(0.25)	(0.29)	(0.55)
No Piped water	-1.60***	-1.30**	-0.83+	-0.22
	(0.45)	(0.42)	(0.49)	(0.94)
State Depriv.				
High Deprivation	0.15	-0.40	-0.17	0.38
	(0.50)	(0.46)	(0.54)	(1.05)
Medium Deprivation	-0.16	-0.44	-0.60	-0.60
	(0.63)	(0.58)	(0.68)	(1.30)
Low Deprivation	0.47	0.33	0.50	0.80
	(0.61)	(0.56)	(0.65)	(1.26)
Very Low Deprivation	0.65	1.36*	1.16+	1.40
	(0.64)	(0.59)	(0.69)	(1.33)
Geo. Region				
Urban-Metrop.	1.36***	1.48***	1.52***	1.80**
	(0.28)	(0.26)	(0.30)	(0.58)
Northeast	-0.34	-0.33	-0.26	-1.19
	(0.43)	(0.40)	(0.46)	(0.90)
West	-1.03*	-0.72+	-0.88+	-2.97**
	(0.44)	(0.41)	(0.47)	(0.91)
Centre	-1.30**	-1.58***	-2.26***	-4.46***
	(0.41)	(0.38)	(0.44)	(0.85)
South	-1.42*	-2.07***	-2.70***	-3.46**
	(0.60)	(0.56)	(0.65)	(1.25)
Southeast	0.75	0.77	-0.11	-1.65
	(0.61)	(0.56)	(0.66)	(1.26)
_cons	85.06***	94.82***	104.41***	121.42***
	(0.92)	(0.85)	(0.99)	(1.91)
N	26,808	26,808	26,808	26,808
r2	.134	.107	.0676	.0224

Notes: standard errors in parenthesis. r2=R squared
+ p<0.1, * p<0.05, ** p<0.01, *** p<0.001
Seg.Pop.=Seguro Popular. PDM=Pemex, Defensa and Marina
Very high deprivation,no health insurance, Northwest region,
piped water in household and rural area used as reference categories

Table A.2.4: Linear regression results for BMI across different percentiles for 2018

	BMI			
	q25	**q50**	**q75**	**q95**
Ethnicity				
Non indigenous	0.16	0.54*	1.07***	2.02**
	(0.25)	(0.24)	(0.32)	(0.69)
Health Affiliation				
IMSS	0.36*	0.21	0.61**	-1.15*
	(0.17)	(0.16)	(0.22)	(0.46)
ISSSTE	0.64*	-0.99***	-0.44	-1.52*
	(0.25)	(0.24)	(0.32)	(0.69)
Seg.Pop.	0.37*	0.20	0.78***	0.40
	(0.17)	(0.16)	(0.21)	(0.45)
PDM	-0.07	1.44*	2.98***	4.41*
	(0.70)	(0.67)	(0.90)	(1.92)
Private	-0.06	-0.26	0.11	-0.72
	(0.40)	(0.38)	(0.51)	(1.09)
Parents' health				
Father not Diabetic	-0.84***	-0.99***	-1.24***	-2.39***
	(0.14)	(0.14)	(0.18)	(0.39)
Father without Hypertension	-0.11	-0.29*	-0.41*	0.25
	(0.13)	(0.13)	(0.17)	(0.37)
Mother not Diabetic.	-0.72***	-0.51***	-0.47**	-0.21
	(0.13)	(0.12)	(0.17)	(0.35)
Mother without Hypertension	-0.54***	-0.48***	-0.42**	-0.46
	(0.12)	(0.11)	(0.15)	(0.33)
Water availability				
Piped water outside	0.15	0.07	-0.07	0.16
	(0.16)	(0.15)	(0.20)	(0.44)
No Piped water	-0.96***	-0.87***	-0.42	0.14
	(0.27)	(0.26)	(0.35)	(0.75)
State Depriv.				
High Deprivation	0.70*	0.91***	0.86*	1.98*
	(0.29)	(0.27)	(0.37)	(0.78)
Medium Deprivation	0.63+	0.76*	0.83+	1.93*
	(0.35)	(0.33)	(0.45)	(0.96)
Low Deprivation	0.92**	1.15***	1.33**	2.17*
	(0.34)	(0.33)	(0.44)	(0.93)
Very Low Deprivation	0.82*	1.69***	1.54***	3.37***
	(0.36)	(0.34)	(0.46)	(0.99)
Geo. Region				
Urban-Metrop.	0.46**	0.44**	0.63**	1.17**
	(0.15)	(0.15)	(0.20)	(0.42)
Northeast	-0.46*	-0.92***	-1.21***	-3.32***
	(0.23)	(0.22)	(0.29)	(0.62)
West	-0.83***	-1.27***	-1.58***	-4.11***
	(0.23)	(0.22)	(0.29)	(0.63)
Centre	-0.69**	-1.19***	-1.74***	-4.40***
	(0.21)	(0.20)	(0.27)	(0.58)
South	0.10	-0.00	-0.46	-2.04*
	(0.33)	(0.32)	(0.42)	(0.91)
Southeast	0.88**	0.80**	0.99*	-0.08
	(0.30)	(0.29)	(0.39)	(0.83)
_cons	25.78***	29.03***	31.97***	39.64***
	(0.49)	(0.47)	(0.63)	(1.34)
N	12,644	12,644	12,644	12,644
r2	.0934	.0853	.0803	.0385

Notes: standard errors in parenthesis. r2=R squared
+ p<0.1, * p<0.05, ** p<0.01, *** p<0.001
Seg.Pop.=Seguro Popular. PDM=Pemex, Defensa and Marina
Very high deprivation,no health insurance, Northwest region,
piped water in household and rural area used as reference categories

Table A.2.5: Linear regression results for WC across different percentiles for 2018

	WC			
	q25	**q50**	**q75**	**q95**
Ethnicity				
Non indigenous	1.35*	2.28***	1.95*	4.85**
	(0.66)	(0.59)	(0.77)	(1.70)
Health Affiliation				
IMSS	0.43	0.82*	1.47**	-0.69
	(0.44)	(0.39)	(0.51)	(1.12)
ISSSTE	-0.06	-0.72	-0.29	-2.75
	(0.66)	(0.59)	(0.77)	(1.70)
Seg.Pop.	-0.02	1.25**	1.81***	1.46
	(0.43)	(0.39)	(0.51)	(1.11)
PDM	-2.89	2.61	9.90***	-6.04
	(1.82)	(1.63)	(2.13)	(4.67)
Private	1.31	1.26	1.59	0.03
	(1.04)	(0.93)	(1.22)	(2.67)
Parents' health				
Father not Diabetic	-3.02***	-2.42***	-3.24***	-5.00***
	(0.37)	(0.33)	(0.44)	(0.96)
Father without Hypertension	-0.02	-0.54+	-0.52	-0.12
	(0.35)	(0.31)	(0.41)	(0.90)
Mother not Diabetic.	-2.06***	-1.17***	-0.88*	0.18
	(0.34)	(0.30)	(0.39)	(0.86)
Mother without Hypertension	-0.80*	-0.73**	0.04	-0.66
	(0.31)	(0.28)	(0.36)	(0.80)
Water availability				
Piped water outside	0.74+	-0.31	-0.09	-2.69*
	(0.42)	(0.38)	(0.49)	(1.08)
No Piped water	-3.22***	-2.15***	-2.07*	-1.46
	(0.72)	(0.65)	(0.84)	(1.85)
State Depriv.				
High Deprivation	5.07***	5.16***	3.92***	1.99
	(0.76)	(0.68)	(0.88)	(1.94)
Medium Deprivation	6.54***	6.39***	4.27***	0.74
	(0.92)	(0.83)	(1.08)	(2.37)
Low Deprivation	7.26***	6.89***	5.04***	2.34
	(0.90)	(0.80)	(1.05)	(2.30)
Very Low Deprivation	6.48***	6.43***	5.57***	3.66
	(0.95)	(0.85)	(1.11)	(2.44)
Geo. Region				
Urban-Metrop.	1.60***	0.82*	1.35**	2.71**
	(0.40)	(0.36)	(0.47)	(1.03)
Northeast	-1.22*	-2.14***	-2.64***	-6.66***
	(0.59)	(0.53)	(0.69)	(1.52)
West	-1.86**	-2.37***	-3.93***	-7.84***
	(0.60)	(0.54)	(0.70)	(1.54)
Centre	-1.15*	-2.30***	-5.52***	-7.78***
	(0.56)	(0.50)	(0.66)	(1.44)
South	4.50***	2.11**	-1.74+	-6.51**
	(0.87)	(0.78)	(1.02)	(2.23)
Southeast	-0.13	-1.09	-3.59***	-8.34***
	(0.80)	(0.72)	(0.94)	(2.05)
_cons	83.96***	91.65***	102.54***	122.92***
	(1.29)	(1.15)	(1.50)	(3.30)
N	12,392	12,392	12,392	12,392
r2	.131	.111	.0742	.0357

Notes: standard errors in parenthesis. r2=R squared
+ $p<0.1$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$
Seg.Pop.=Seguro Popular. PDM=Pemex, Defensa and Marina
Very high deprivation,no health insurance, Northwest region,
piped water in household and rural area used as reference categories

A.3 Shapley decomposition

Table A.3.1: Relative contribution of each circumstance to IOp in BMI and WC across time

Outcome/year	Ethnicity	Health Ins	Parent's health	Avail. Water	State Depriv.	Urbanity	Geo. Region
BMI 2012	2.70	3.75	64.26	3.16	4.14	8.22	13.77
BMI 2018	1.27	4.97	59.98	3.14	5.88	4.98	19.79
WC 2012	4.71	8.38	57.01	5.48	7.41	7.68	9.34
WC 2018	2.24	6.05	57.96	5.51	11.38	3.51	13.35

Table A.3.2: Relative contribution of each circumstance to IOp in different percentiles of BMI and WC across time

Outcome/year	Pth	Ethnicity	Health Ins	Parent health	Avail. Water	State Depriv.	Urbanity	Geo Region
BMI 2012	25	0.85	15.37	55.08	5.04	4.24	6.51	12.91
	50	3.60	5.53	60.22	4.57	3.88	10.00	12.21
	75	6.87	1.14	57.56	2.37	6.31	8.26	17.49
	95	4.75	3.81	61.46	0.93	2.87	9.11	17.07
BMI 2018	25	0.20	10.27	66.89	4.35	2.37	3.17	12.76
	50	1.18	8.04	51.13	4.09	11.27	4.06	20.23
	75	2.77	12.74	44.14	3.35	7.59	5.03	24.39
	95	3.48	11.60	25.95	0.72	12.71	5.03	40.52
WC 2012	25	2.29	22.31	54.19	5.66	4.18	5.75	5.63
	50	6.18	6.07	50.87	6.86	11.76	7.91	10.34
	75	8.92	3.47	49.60	5.05	11.13	7.73	14.10
	95	9.70	6.97	48.90	2.20	7.15	7.17	17.91
WC 2018	25	0.95	6.73	60.86	6.68	10.58	4.66	9.54
	50	4.45	8.24	49.30	5.82	17.54	2.63	12.01
	75	2.86	16.52	30.93	4.81	14.37	3.57	26.94
	95	7.06	5.46	26.09	8.69	14.05	8.50	30.15

A.4 IOp results using the D-index

The dissimilarity index is an absolute measure of IOp. It measures disparities on the level of health across types compared with the average level of health of the population. The index ranges between 0 and 1, if equality exists, D=0. (Paes de Barros et al., 2008).

$$\theta_a = D(\hat{y}) = \frac{1}{2N\bar{\hat{y}}} \sum_{i=1}^{N} |\hat{y}_i - \bar{\hat{y}}|$$

where $\hat{y_i} = \mathbb{E}(y|C_i)$. Since y is a binary variable, the model used to estimate y was a logit model.

Table A.4.1: Dissimilarity Index for BMI and WC in 2012 and 2018

Outcome-Year	***DI***	***BSE***
BMI-2012	0.0708***	0.0089
BMI-2018	0.0564***	0.0090
WC-2012	0.0881***	0.0064
WC-2018	0.0670***	0.0096

Notes: BSE=bootstrapped standard errors (500 replications)

BMI and WC dichotomised for excess weight and excess adiposity

Excess weight: 1 if BMI>25kg/m2, 0 otherwise.

Excess adiposity: 1 if WC > 80 cm, 0 otherwise.

+ p<0.1, * p<0.05, ** p<0.01, *** p<0.001

Appendix B

Chapter 2

B.1 Variable definitions

Table B.1.1: Key variables definition

Outcomes		
Variable	**Definition**	**Focus**
Stunting	HAZ <-2 z scores	Under nutrition
Wasting	WHZ <-2 z scores	
Underweight	WAZ <-2 z scores	
Overweight	BMIZ >2 z scores	Over nutrition
Double burden in children	HAZ or WHZ or WAZ <-2 z scores AND BMI >2 z scores	Malnutrition
Anaemia*	HB<13 gl/dl	Under nutrition
Excess weight*	BMI>25 kg/mts2	Over nutrition
Excess adiposity*	WC >80 cm	
Double burden in people older than 11 years old	BMI >25 kg/mts2 or WC>80cm AND HB<13 gl/dl	Malnutrition
Circumstances		
Variable	**Definition**	
Sex	1 if man, 0 woman	
Ethnicity	1 if non-indigenous, 0 if indigenous	
Mother's health insurance	1 if mother insuranced, 0 otherwise	
Mother's BMI	In kg/mts2	
Mother's anaemia	1 if mother anaemic, 0 otherwise	
Parental diabetes	1 if mother or father was diagnosed with diabetes, 0 otherwise	
Parent's education	Five categories: No education, up to primary school up to secondary school, up to high school and higher education	
Running water available in the household	1 if water available, 0 otherwise	
Household living standards	Five levels: low, medium low, medium high and high	
State deprivation level	Five levels: very high,high,medium,low and very low	
Geographic region	Six regions: Northwest, Northeast, West, Centre, South and Southeast	
Efforts		
Dietary patterns	Three patterns: Low-nutritious and highly-caloric, high-nutritious food and legume and maize-based products. Food supplements consumption	
Physical activity	Daily hours dedicated to do vigorous and moderate physical activities	
Alcohol consumption	1 if the alcohol consumption is above recommendation, 0 otherwise	
Smoking frequency	1 if smoking tobacco occurs daily or regularly, 0 otherwise	

Note: *Outcomes measured in individuals older than 11 years old

Table B.1.2: Outcomes analysed in the *Ex-ante* and *Ex-post* approaches

Ex-ante approach						**Ex-post approach**				
		Mean-based		**Beyond-the-mean**			**Mean-based**		**Beyond-the-mean**	
Age group	**Outcome**	**Proxy**	**Focus**	**Outcome**	**Focus**	**Outcome**	**Proxy**	**Focus**	**Outcome**	**Focus**
0 to 5 years old	Stunting	HAZ <-2 z scores	Under nutrition							
	Wasting	WHZ <-2 z scores								
	Underweight	WAZ <-2 z scores								
	Overweight	BMIZ >2 z scores	Over nutrition							
	Double burden	HAZ or WHZ or WAZ <-2 z scores AND BMI >2 z scores	Malnutrition							
11 to 35 years old	Anaemia	HB<13 gl/dl	Under nutrition	HB		Anaemia	HB<13 gl/dl	Under nutrition	HB	
	Excess weight	BMI>25 kg/mts2	Over nutrition	BMI	Percentiles: 10^{th} 25^{th} 50^{th} 75^{th} 90^{th} 95^{th} and 99^{th}	Excess weight	BMI>25 kg/mts2	Over nutrition	BMI	Percentiles: 10^{th} 25^{th} 50^{th} 75^{th} 90^{th} 95^{th} and 99^{th}
	Excess adiposity	WC >80 cm		WC		Excess adiposity	WC >80 cm		WC	
	Double burden	BMI >25 kg/mts2 or WC>80cm AND HB<13 gl/dl	Malnutrition			Double burden	BMI >25 kg/mts2 or WC>80cm AND HB<13 gl/dl	Malnutrition		

B.2 Matching results

Table B.2.1: Matching summary

		Control	**Treatment**	**Strata**
1988-1999	All	6,471	6,510	60
	Matched	6,471	6,510	60
	Unmatched	0	0	0
1988-2006	All	6,517	6,510	65
	Matched	2,201	6,455	59
	Unmatched	4,316	55	6
1988-2012	All	14,559	6,510	60
	Matched	14,559	6,510	60
	Unmatched	0	0	0
1988-2016	All	2,223	6,510	60
	Matched	2,223	6,510	60
	Unmatched	0	0	0
1988-2018	All	2,208	6,510	60
	Matched	2,208	6,510	60
	Unmatched	0	0	0

The 55 treatment observations unmatched across the 1988 and 2006 surveys correspond to men that lived in the Northwest region and were born in 1985.

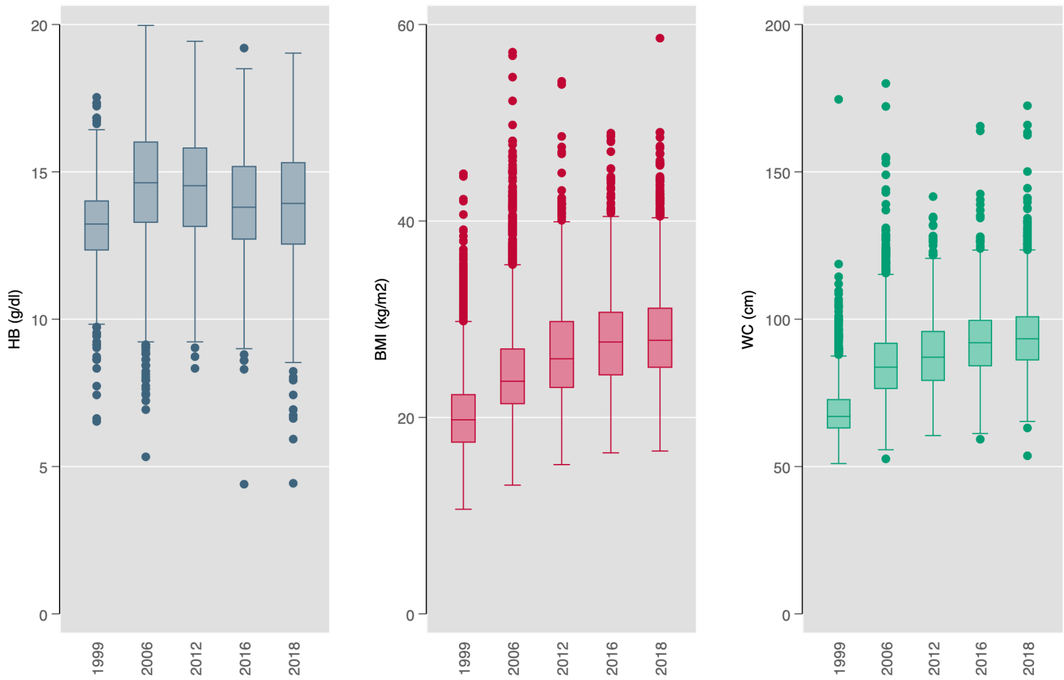
B.3 Distribution of samples by year of birth

Table B.3.1: Distribution of the matched samples across cohorts

	Survey year						
Year of birth	1988	1999	2006	2012	2016	2018	**Total**
1983		949	694	1,565	406	372	3,986
1984	1,212	1,063	706	1,589	400	384	5,354
1985	1,373	1,071	677	1,661	334	363	5,479
1986	1,295	1,161	716	1,704	377	382	5,635
1987	1,258	1,514	1,336	1,750	364	331	6,553
1988	1,353	1,337	1,531	1,719	342	376	6,658
Total	6,491	7,095	5,660	9,988	2,223	2,208	33,665

B.4 Description of continuous outcomes

Figure B.4.1: Distribution of HB, BMI and WC across survey years



B.5 Construction of effort variables

Eating pattern 1: low-nutritious and high-energy food, such as whole-fat diary, fast food, sweetened beverages, sweets and red meat. **Eating pattern 2:** high-nutritious food, such as fruits, vegetables, poultry, fish, and cereals. **Eating pattern 3:** Legumes and maize-based products.

Figure B.5.1: Distribution of eating-patterns across survey years

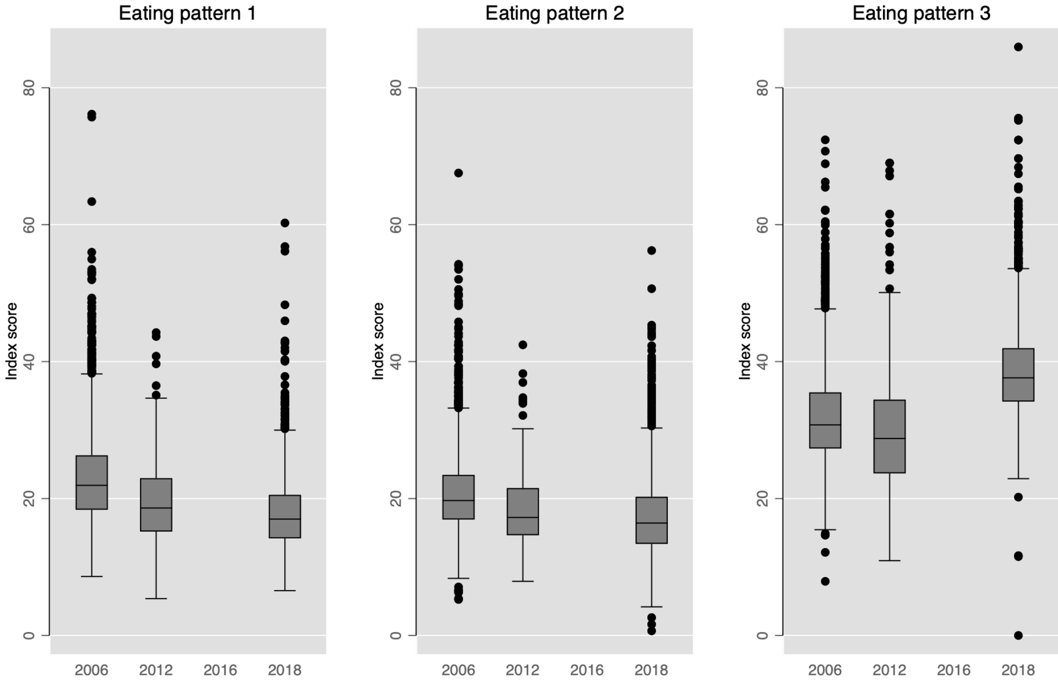


Table B.5.1: Food items included in the 2006, 2012 and 2018 surveys (I)

Animal protein	**One portion is:**
Milk (whole, skim, lactose-free, powder, LICONSA, etc)	1 cup (240 ml)
Panela or cottage-type cheese	1 slice or 2 tablespoons (30 g)
Chihuahua-type, manchego, gouda cheese	1 slice (30 g)
Yogurt	1 typical cup yogurt (150 g)
Liquid yoghurt	One typical jar (230 g)
Petite suisse yogurt	1 pot (45 g)
Sweetened probiotic milk beverage	1 bottle (80 ml)
Pork or beef meat	1 medium steak (90 g)
Dried beef	1 plate (80 g)
Sausage	1/2 piece (30 g)
Pork sausage, ham, turkey, mortadella	1 sausage or slice (30 g)
Chicken (leg, thigh or breast)	1 piece (leg, thigh) (90 g)
Chicken wing	One wing
Chiken liver or gizzard	1 piece (30 g)
Egg (boiled or cooked)	1 piece (62 g)
Fish	1 piece (90 g)
Tuna and sardines (Tomato, water or oil)	1/4 tin (40 g)
Other seafood (shrimp, oysters, etc.)	1 plate (100 g)
Fruits, vegetables and legumes	**One portion is:**
Banana	1 medium piece (176 g)
Jicama	1/2 medium piece (163 g)
Orange or tangerine	1 large piece (206 g)
Apple or pear	1 medium piece (140 g)
Melon or watermelon	1 slice or 3/4 cup (115 g)
Guava	1 medium piece (75 g)
Mango	1 medium piece (185 g)
Papaya	1 medium slice (100 g) or 1/2 cup
Pineapple	1 medium slice (150 g)
Grapefruit	1 piece small (270 g)
Strawberry	1 cup or 9 medium pieces (140 g)
Grapes	10 pieces (60 g)
Peach	1 medium piece (50 g)
Red tomatoes	Half piece (30 g)
Green leaves (spinach, quelites)	1/2 plate (85 g)
Squash	1/4 piece(50 g) or 1/3 cup
Carrot	1 small piece or 1/2 cup (50 g)
Zucchini	1/2 medium part (50 g)
Broccoli, cauliflower or cabbage	1/4 cup (35 g)
Green beans	1/4 cup or 5 pieces (30 g)
Corn cob	Half small piece (50 g)
Lettuce	1/2 cup (30 g)
Nopales	1 large piece (100 g)
Cucumber	1/2 large piece (150 g)
Avocado	1 slice or 1 small piece (33 g)
Poblano chile	A medium piece or 1/3 cup (80 g)
Canned vegetables such as peas, carrots, mushrooms and green beans	1/3 cup or 1 small can
Frozen vegetables such as peas, carrots, broccoli, cauliflower, green beans	1/3 cup
Beans	50 g
Lentils and chick peas	1 plate or 1 cup (100 g)

Table B.5.2: Food items included in the 2006, 2012 and 2018 surveys (II)

Energy-dense food	**One portion is:**
Cake or sandwich	1 medium piece (130 g)
Burger	Medium 1 piece (240 g)
Pizza	1 slice small (92 g)
Hot dog	1 medium piece (110 g)
Soda	1 cup (240 ml)
Tea or coffee	1 cup (240 ml)
Natural juices	1 cup (240 ml)
Fruit-favoured water	1 cup (240 ml)
Sweetened beverages	1 cup (240 ml)
Chocolate	1 piece or 1 tablespoon (10 g)
Candy (Candies, lollipops)	1 piece (30 g)
Spicy and sweet candy	1 piece (30 g)
Snacks (like peanuts, crisps)	1 single package or small bag (35 g)
Gelatin, flan	1 piece or slice (125 g)
Cake or pie	1 medium slice (125 g)
Ice cream	1 scoop (80 g)
Peanuts, beans or seeds	1 fist (han (35 g)
Microwavable popcorn (All types)	1 medium bag (100 g)
Donuts	1 piece (70g)
Biscuits (All types)	2 units (32 g)
Cereal bars	1 piece (25g)
Margarine	1 tablespoon (10 g)
Butter	1 tablespoon (10 g)
Mayonnaise	1 tablespoon (10 g)
Fresh sour cream	1 tablespoon (10 g)
Vegetable lard	1 tablespoon (10 g)
Animal lard	1 tablespoon (10 g)
Pasta, cereals and white bread	**One portion is:**
Cooked rice	1 cup or 1 plate (100 g)
White bread	2 slices or one roll(70g)
Wholemeal bread	2 slices or one roll (70g)
Sweet bread (Except donuts and fritters)	1 piece (70g)
Bakery donuts and churros	1 piece (70g)
Pretzels	4 pieces (20 g)
Potatoes cooked	1/2 medium cooked piece (40g)
Fried potatos	1/2 medium piece (40 g)
Box cereals (chocolate, light, corn flakes, fruit-flavoured, fiber)	1 cup (30 g dry)
Broth	1 cup (240 ml)
Soups	1 plate (240 mL)
Instant soup	1 pot (64 g)
Maize products and tortillas	**One portion is:**
Meatless snacks like sopes, quesadillas, tlacoyos, gorditas and enchiladas	100 g
Antojitos with beef, pork, poultry, organ meats, such as tacos, quesadillas, tlacoyos, enchiladas, gorditas	100 g
Pozole (All types)	1 plate (100 g)
Tamale (All types)	1 piece (200 g)
Atole (all types)	1 cup (240 ml)
Maize tortillas	1 unit
Wheat flour tortilla	1 unit

B.6 Nonlinear and linear regression models. *Ex-ante* approach

Table B.6.1: Logit regression results: *Ex-Ante* analysis, 1988. Clinical cut-off points.

	HAZ	**WHZ**	**WAZ**	**BMIz**	**DBM**
Individual Characteristics					
Individual's sex	0.078	-0.004	-0.002	0.295**	0.265+
	(0.068)	(0.110)	(0.088)	(0.100)	(0.156)
Individual's ethnicity	-0.616***	-0.155	-0.768***	0.561*	0.392
	(0.135)	(0.274)	(0.148)	(0.282)	(0.331)
Household living standards					
Running water HH	0.183+	0.419**	0.363**	-0.088	0.224
	(0.100)	(0.159)	(0.132)	(0.151)	(0.243)
LS:Low	2.460***	0.357	2.011***	0.129	1.551***
	(0.161)	(0.229)	(0.216)	(0.200)	(0.339)
LS:Med-Low	1.326***	0.407*	1.317***	-0.422*	0.160
	(0.149)	(0.189)	(0.200)	(0.182)	(0.332)
LS:Med-High	0.876***	0.396*	0.900***	-0.192	0.191
	(0.137)	(0.163)	(0.187)	(0.146)	(0.287)
State Deprivation					
Dep:V.High	0.129	0.146	0.160	0.259	-0.167
	(0.155)	(0.307)	(0.205)	(0.245)	(0.404)
Dep:High	0.198	-0.050	-0.199	-0.143	0.190
	(0.148)	(0.210)	(0.192)	(0.204)	(0.330)
Dep:Med	0.312	-0.568	0.049	-0.354	-0.130
	(0.204)	(0.368)	(0.275)	(0.280)	(0.481)
Dep:Low	0.099	-0.103	-0.042	0.003	0.106
	(0.105)	(0.158)	(0.143)	(0.142)	(0.255)
Geographical Region					
Northwest	-0.689**	0.930+	-1.027**	0.780*	0.496
	(0.264)	(0.518)	(0.346)	(0.362)	(0.603)
Northeast	-0.172	1.079*	-0.229	0.148	0.493
	(0.206)	(0.478)	(0.259)	(0.336)	(0.509)
West	0.217	1.069*	-0.111	0.615+	0.988*
	(0.202)	(0.477)	(0.256)	(0.327)	(0.494)
Centre	0.264	0.168	-0.489+	-0.166	0.364
	(0.217)	(0.497)	(0.278)	(0.351)	(0.539)
South	0.157	0.385	-0.268	-0.475	0.639
	(0.258)	(0.581)	(0.325)	(0.427)	(0.645)
Constant	-2.412***	-3.751***	-2.597***	-3.146***	-5.305***
	(0.302)	(0.604)	(0.378)	(0.490)	(0.720)
N	5,764	5,837	5,981	5,861	5,475
r2_p	.129	.0251	.0748	.0224	.052

Notes: standard errors in parenthesis + p<0.1, * p<0.05, ** p<0.01, *** p<0.001
r2_p=pseudo R squared. HAZ=1 if stunting. WHZ=1 if wasting. WAZ=1 if underweight
BMIz=1 if overweight. DBM=1 if stunting or wasting or underweight and overweight

Table B.6.2: Logit regression results: *Ex-Ante* analysis, 1999. Clinical cut-off points.

	DBM	**HB**	**BMI**	**WC**
Individual Characteristics				
Individual's sex	0.000	1.323***	-1.711***	0.000
	(.)	(0.253)	(0.443)	(.)
Individual's ethnicity	-0.415	-0.256	0.912+	0.894
	(1.167)	(0.345)	(0.469)	(0.688)
Household living standards				
Running water HH	-0.102	0.238	-0.121	-0.346
	(0.903)	(0.293)	(0.210)	(0.250)
LS:Low	-2.280+	0.394	-0.804**	-1.156**
	(1.259)	(0.427)	(0.297)	(0.364)
LS:Med-Low	-1.007	0.590	-0.756**	-0.904**
	(0.959)	(0.367)	(0.249)	(0.294)
LS:Med-High	-2.070*	-0.287	-0.351+	-0.592**
	(0.978)	(0.336)	(0.188)	(0.227)
State Deprivation				
Dep:V.High	0.487	-0.314	-0.569	-1.243
	(2.438)	(0.701)	(0.576)	(0.770)
Dep:High	0.813	0.485	0.011	-0.482
	(1.104)	(0.411)	(0.274)	(0.339)
Dep:Med	0.035	0.636	-0.061	-0.094
	(1.101)	(0.460)	(0.295)	(0.338)
Dep:Low	0.010	0.694*	0.053	-0.191
	(0.968)	(0.353)	(0.215)	(0.250)
Geographical Region				
Northwest	1.836	-0.159	0.652	0.234
	(2.214)	(0.639)	(0.520)	(0.697)
Northeast	-0.064	-0.344	0.338	0.176
	(2.260)	(0.597)	(0.506)	(0.678)
West	0.961	-0.822	0.124	0.238
	(2.148)	(0.588)	(0.503)	(0.670)
Centre	0.069	-0.903	-0.386	-0.459
	(2.111)	(0.555)	(0.484)	(0.655)
South	0.000	0.000	0.000	0.000
	(.)	(.)	(.)	(.)
Constant	-2.677	-1.594+	-2.132**	-1.785+
	(2.730)	(0.822)	(0.731)	(1.003)
N	518	775	1,905	1,519
r2_p	.107	.0583	.0676	.0647

Notes: standard errors in parenthesis + $p<0.1$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$
r2_p=pseudo R squared. HB=1 if anaemia. BMI=1 if excess weight
WC=1 if excess adiposity. DBM=1 if excess weight or adiposity and anaemia
Sex and South omitted because of collinearity.

Table B.6.3: Logit regression results: *Ex-Ante* analysis, 2006. Clinical cut-off points.

	DBM	**HB**	**BMI**	**WC**
Individual Characteristics				
Individual's sex	-1.534***	-1.245***	0.082	-0.918***
	(0.161)	(0.107)	(0.058)	(0.084)
Individual's ethnicity	0.170	0.370	0.414*	0.637**
	(0.361)	(0.283)	(0.185)	(0.234)
Household living standards				
Running water HH	0.016	0.056	-0.030	0.029
	(0.185)	(0.130)	(0.082)	(0.112)
LS:Low	0.479*	0.253	-0.061	0.162
	(0.223)	(0.158)	(0.098)	(0.136)
LS:Med-Low	0.344+	0.073	0.130	0.195+
	(0.200)	(0.142)	(0.083)	(0.118)
LS:Med-High	-0.062	0.021	-0.081	0.017
	(0.206)	(0.141)	(0.080)	(0.115)
State Deprivation				
Dep:V.High	-0.421	0.291	-0.720***	-0.712*
	(0.483)	(0.333)	(0.216)	(0.311)
Dep:High	-0.075	0.558**	-0.309**	-0.300*
	(0.258)	(0.185)	(0.108)	(0.147)
Dep:Med	-0.164	0.113	-0.134	-0.030
	(0.234)	(0.177)	(0.100)	(0.139)
Dep:Low	-0.057	0.190	0.057	0.138
	(0.227)	(0.170)	(0.093)	(0.128)
Geographical Region				
Northwest	0.292	0.636+	-0.526**	-0.133
	(0.477)	(0.337)	(0.202)	(0.298)
Northeast	0.082	0.338	-0.507**	-0.038
	(0.436)	(0.305)	(0.182)	(0.277)
West	0.049	0.295	-0.524**	-0.304
	(0.446)	(0.312)	(0.187)	(0.284)
Centre	-0.558	-0.132	-0.464**	-0.369
	(0.421)	(0.293)	(0.174)	(0.266)
South	-0.005	0.185	-0.208	0.140
	(0.524)	(0.361)	(0.227)	(0.350)
Constant	-1.978***	-2.534***	-0.304	-0.051
	(0.580)	(0.432)	(0.264)	(0.362)
N	3,005	4,863	5,100	2,430
r2_p	.0815	.0596	.00881	.0471

Notes: standard errors in parenthesis + p<0.1, * p<0.05, ** p<0.01, *** p<0.001
r2_p=pseudo R squared. HB=1 if anaemia. BMI=1 if excess weight
WC=1 if excess adiposity. DBM=1 if excess weight or adiposity and anaemia

Table B.6.4: Logit regression results: *Ex-Ante* analysis, 2012. Clinical cut-off points.

	DBM	**HB**	**BMI**	**WC**
Individual Characteristics				
Individual's sex	-1.777***	-1.476***	0.015	-0.993***
	(0.433)	(0.313)	(0.130)	(0.138)
Individual's ethnicity	1.306	0.932+	0.076	0.104
	(0.840)	(0.505)	(0.251)	(0.269)
Household living standards				
Running water HH	-0.765	-0.584+	0.134	0.185
	(0.466)	(0.318)	(0.182)	(0.195)
LS:Low	-0.716	-0.263	0.092	-0.245
	(0.581)	(0.442)	(0.244)	(0.260)
LS:Med-Low	-0.349	-0.197	0.614**	0.517*
	(0.540)	(0.400)	(0.217)	(0.232)
LS:Med-High	-0.458	-0.232	0.257	-0.020
	(0.496)	(0.376)	(0.194)	(0.204)
State Deprivation				
Dep:V.High	1.928	2.661*	0.726	1.128*
	(1.484)	(1.325)	(0.452)	(0.476)
Dep:High	0.672	0.774+	-0.375	0.027
	(0.648)	(0.466)	(0.244)	(0.254)
Dep:Med	0.461	0.335	-0.179	0.122
	(0.581)	(0.421)	(0.215)	(0.228)
Dep:Low	0.820	0.719+	-0.012	0.344
	(0.588)	(0.408)	(0.217)	(0.228)
Geographical Region				
Northwest	-1.214	-0.589	-0.481	-0.022
	(0.956)	(0.679)	(0.460)	(0.472)
Northeast	-1.351+	-0.789	-0.885*	-0.258
	(0.801)	(0.584)	(0.407)	(0.419)
West	-1.657+	-1.077+	-0.932*	-0.411
	(0.860)	(0.632)	(0.424)	(0.435)
Centre	-1.543*	-1.133*	-0.501	-0.135
	(0.784)	(0.567)	(0.395)	(0.405)
South	-2.577+	-2.676*	-1.079*	-0.616
	(1.412)	(1.298)	(0.491)	(0.508)
Constant	-1.155	-1.849*	0.697	0.594
	(1.331)	(0.919)	(0.549)	(0.574)
N	498	1,049	1,122	1,051
r2_p	.125	.0873	.0241	.0652

Notes: standard errors in parenthesis + p<0.1, * p<0.05, ** p<0.01, *** p<0.001
r2_p=pseudo R squared. HB=1 if anaemia. BMI=1 if excess weight
WC=1 if excess adiposity. DBM=1 if excess weight or adiposity and anaemia

Table B.6.5: Logit regression results: *Ex-Ante* analysis, 2018. Clinical cut-off points.

	DBM	**HB**	**BMI**	**WC**
Individual Characteristics				
Individual's sex	-1.974***	-1.648***	0.034	-0.998***
	(0.170)	(0.130)	(0.102)	(0.114)
Individual's ethnicity	-0.340	-0.330	0.079	0.091
	(0.310)	(0.218)	(0.198)	(0.194)
Household living standards				
Running water HH	-0.068	-0.027	-0.144	-0.404*
	(0.233)	(0.168)	(0.150)	(0.159)
LS:Low	0.139	-0.293	0.490*	0.795***
	(0.322)	(0.235)	(0.200)	(0.213)
LS:Med-Low	0.293	0.088	0.433*	0.772***
	(0.286)	(0.206)	(0.186)	(0.200)
LS:Med-High	0.194	-0.092	0.372*	0.514**
	(0.248)	(0.179)	(0.159)	(0.168)
State Deprivation				
Dep:V.High	0.163	-0.018	0.670+	0.073
	(0.596)	(0.453)	(0.401)	(0.406)
Dep:High	0.630+	0.509+	0.461*	0.308
	(0.354)	(0.273)	(0.212)	(0.225)
Dep:Med	0.591+	0.396	0.472*	0.684**
	(0.323)	(0.253)	(0.195)	(0.212)
Dep:Low	0.979**	0.791**	0.267	0.506*
	(0.341)	(0.263)	(0.204)	(0.220)
Geographical Region				
Northwest	-0.506	0.164	-0.903*	-0.339
	(0.573)	(0.393)	(0.424)	(0.386)
Northeast	-0.320	0.030	-0.715+	-0.082
	(0.527)	(0.348)	(0.391)	(0.340)
West	-1.140*	-0.555	-0.914*	-0.373
	(0.551)	(0.374)	(0.402)	(0.353)
Centre	-0.565	-0.107	-0.924*	-0.255
	(0.501)	(0.334)	(0.379)	(0.322)
South	-0.118	0.164	-0.720	-0.087
	(0.636)	(0.450)	(0.477)	(0.434)
Constant	0.565	-0.716	1.258**	1.306**
	(0.645)	(0.455)	(0.458)	(0.420)
N	827	2,078	2,112	2,019
r2_p	.161	.104	.0123	.0564

Notes: standard errors in parenthesis + $p<0.1$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$
r2_p=pseudo R squared. HB=1 if anaemia. BMI=1 if excess weight
WC=1 if excess adiposity. DBM=1 if excess weight or adiposity and anaemia

Table B.6.6: Linear regression: *Ex-Ante* IOp in BMI across different percentiles. 1999

	p10	**p25**	**p50**	**p75**	**p90**	**p95**	**p99**
Individual Characteristics							
Individual's sex	-1.040***	-2.838***	-3.191***	-2.512***	-2.913***	-3.487**	-4.744
	(0.238)	(0.285)	(0.340)	(0.476)	(0.747)	(1.200)	(3.169)
Individual's ethnicity	0.451	0.520	0.493	1.155*	0.743	0.980	-1.695
	(0.287)	(0.344)	(0.411)	(0.574)	(0.902)	(1.450)	(3.827)
Household living standards							
Running water HH	0.425+	0.051	-0.476	0.265	-0.217	-0.926	-3.814
	(0.220)	(0.264)	(0.315)	(0.440)	(0.691)	(1.111)	(2.933)
LS:Low	-0.109	-1.101**	-2.313***	-2.423***	-3.134**	-4.309**	-9.420*
	(0.309)	(0.370)	(0.442)	(0.618)	(0.970)	(1.558)	(4.115)
LS:Med-Low	0.048	-0.742*	-1.673***	-2.215***	-2.897***	-4.379**	-7.021*
	(0.267)	(0.320)	(0.382)	(0.534)	(0.839)	(1.348)	(3.560)
LS:Med-High	-0.235	-0.418	-1.252***	-1.398**	-1.886**	-2.466*	-7.893**
	(0.219)	(0.262)	(0.313)	(0.438)	(0.687)	(1.104)	(2.916)
State Deprivation							
Dep:V.High	-0.761	-0.765	-0.903	0.005	-2.325	-2.210	-2.344
	(0.569)	(0.682)	(0.815)	(1.139)	(1.789)	(2.874)	(7.588)
Dep:High	-0.547+	-0.526	-0.500	0.190	-0.622	-0.359	-5.186
	(0.307)	(0.367)	(0.439)	(0.614)	(0.964)	(1.548)	(4.088)
Dep:Med	-0.409	-0.709+	-0.199	-0.052	-0.543	-0.860	-1.334
	(0.338)	(0.405)	(0.484)	(0.676)	(1.062)	(1.706)	(4.505)
Dep:Low	-0.287	-0.598+	-0.185	-0.044	-0.041	-0.815	-2.388
	(0.259)	(0.310)	(0.371)	(0.519)	(0.815)	(1.309)	(3.456)
Geographical Region							
Northwest	-0.561	0.057	0.539	1.655	1.764	1.175	2.298
	(0.533)	(0.639)	(0.764)	(1.067)	(1.676)	(2.694)	(7.112)
Northeast	-0.847+	-0.523	-0.551	0.589	0.799	0.027	7.461
	(0.512)	(0.613)	(0.734)	(1.025)	(1.610)	(2.586)	(6.829)
West	-0.775	-0.545	-0.169	0.213	-0.205	0.041	4.233
	(0.503)	(0.602)	(0.720)	(1.006)	(1.580)	(2.538)	(6.702)
Centre	-0.652	-0.559	-0.409	-0.498	-1.692	-2.243	0.350
	(0.472)	(0.566)	(0.676)	(0.945)	(1.484)	(2.384)	(6.296)
South	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	(.)	(.)	(.)	(.)	(.)	(.)	(.)
Constant	17.052***	19.356***	22.293***	23.448***	28.768***	32.884***	45.797***
	(0.654)	(0.783)	(0.936)	(1.308)	(2.054)	(3.301)	(8.716)
N	1,931	1,931	1,931	1,931	1,931	1,931	1,931
r2	.0226	.0732	.0799	.0591	.0497	.0223	.0153

Notes: standard errors in parenthesis + p<0.1, * p<0.05, ** p<0.01, *** p<0.001
r2=R squared. South region omitted because of collinearity.

Table B.6.7: Linear regression: *Ex-Ante* IOp in WC across different percentiles. 1999

	p10	p25	p50	p75	p90	p95	p99
Individual Characteristics							
Individual's sex	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	(.)	(.)	(.)	(.)	(.)	(.)	(.)
Individual's ethnicity	0.778	-0.638	0.121	2.940+	2.192	-1.576	-0.235
	(0.989)	(0.951)	(0.990)	(1.564)	(2.394)	(4.527)	(8.428)
Household living standards							
Running water HH	0.723	0.192	0.468	-0.474	-1.791	-6.338+	-7.626
	(0.712)	(0.685)	(0.713)	(1.126)	(1.723)	(3.258)	(6.066)
LS:Low	-1.454	-2.704**	-2.868**	-6.519***	-8.051***	-16.313***	-12.221
	(0.997)	(0.959)	(0.998)	(1.577)	(2.414)	(4.564)	(8.497)
LS:Med-Low	-0.167	-2.062*	-2.785**	-5.334***	-6.971***	-11.505**	-8.092
	(0.856)	(0.823)	(0.857)	(1.354)	(2.073)	(3.920)	(7.297)
LS:Med-High	-0.798	-1.496*	-2.579***	-4.656***	-4.943**	-8.371*	-3.648
	(0.713)	(0.685)	(0.713)	(1.127)	(1.725)	(3.261)	(6.071)
State Deprivation							
Dep:V.High	0.221	0.507	1.222	0.294	-4.897	-19.518*	-10.954
	(1.895)	(1.822)	(1.897)	(2.996)	(4.586)	(8.672)	(16.144)
Dep:High	0.214	0.485	0.568	-1.491	-2.338	-12.422**	-12.805
	(0.988)	(0.950)	(0.990)	(1.563)	(2.393)	(4.524)	(8.423)
Dep:Med	1.756	1.797+	0.670	-0.731	-0.664	-9.236+	-0.641
	(1.100)	(1.058)	(1.102)	(1.740)	(2.663)	(5.037)	(9.376)
Dep:Low	0.537	1.237	0.931	0.093	-1.581	-8.228*	-0.350
	(0.825)	(0.793)	(0.826)	(1.304)	(1.996)	(3.775)	(7.027)
Geographical Region							
Northwest	0.820	0.684	3.271+	6.067*	3.454	2.357	2.188
	(1.806)	(1.737)	(1.808)	(2.856)	(4.372)	(8.267)	(15.389)
Northeast	1.148	1.592	3.938*	4.757+	3.180	-1.891	1.292
	(1.721)	(1.655)	(1.723)	(2.722)	(4.167)	(7.879)	(14.668)
West	-0.655	-0.684	1.719	3.597	3.091	-0.361	3.389
	(1.706)	(1.640)	(1.708)	(2.697)	(4.129)	(7.808)	(14.535)
Centre	0.086	0.643	2.277	2.320	-1.453	-7.575	-0.839
	(1.603)	(1.541)	(1.604)	(2.534)	(3.879)	(7.335)	(13.655)
South	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	(.)	(.)	(.)	(.)	(.)	(.)	(.)
Constant	58.522***	64.010***	66.683***	72.712***	86.099***	113.813***	115.856***
	(2.189)	(2.105)	(2.192)	(3.462)	(5.300)	(10.022)	(18.657)
N	1,505	1,505	1,505	1,505	1,505	1,505	1,505
r2	.0181	.0247	.0407	.0589	.0413	.0392	.00991

Notes: standard errors in parenthesis + p<0.1, * p<0.05, ** p<0.01, *** p<0.001
r2=R squared. South region omitted because of collinearity.

Table B.6.8: Linear regression: *Ex-Ante* IOp in BMI across different percentiles. 2006

	p10	**p25**	**p50**	**p75**	**p90**	**p95**	**p99**
Individual Characteristics							
Individual's sex	0.266+	0.278*	0.256+	0.121	-0.320	-1.025*	-1.247
	(0.137)	(0.128)	(0.149)	(0.209)	(0.371)	(0.436)	(1.143)
Individual's ethnicity	-0.546	0.367	0.825+	1.447*	3.132**	1.790	2.162
	(0.396)	(0.370)	(0.431)	(0.605)	(1.075)	(1.264)	(3.312)
Household living standards							
Running water HH	0.015	-0.153	-0.020	0.029	-0.451	-1.402*	-2.220
	(0.192)	(0.180)	(0.209)	(0.294)	(0.521)	(0.613)	(1.606)
LS:Low	0.221	0.035	0.099	-0.523	0.091	-0.042	1.924
	(0.229)	(0.214)	(0.249)	(0.350)	(0.622)	(0.731)	(1.917)
LS:Med-Low	0.203	0.363*	0.443*	0.245	-0.056	0.514	3.184+
	(0.196)	(0.183)	(0.213)	(0.300)	(0.532)	(0.626)	(1.640)
LS:Med-High	0.125	0.191	-0.007	-0.509+	-1.077*	-1.148+	1.205
	(0.190)	(0.177)	(0.206)	(0.290)	(0.515)	(0.606)	(1.587)
State Deprivation							
Dep:V.High	0.029	-0.741	-2.465***	-1.879*	-2.215	-1.041	-2.672
	(0.503)	(0.471)	(0.547)	(0.769)	(1.366)	(1.606)	(4.209)
Dep:High	0.438+	0.046	-0.845**	-1.141**	-2.121**	-1.802*	-4.319*
	(0.252)	(0.236)	(0.274)	(0.386)	(0.685)	(0.806)	(2.111)
Dep:Med	0.106	0.061	-0.271	-0.824*	-1.545*	-1.750*	-3.597+
	(0.238)	(0.222)	(0.258)	(0.363)	(0.645)	(0.759)	(1.988)
Dep:Low	0.018	0.074	-0.169	-0.178	-0.627	-0.571	-1.178
	(0.221)	(0.207)	(0.241)	(0.339)	(0.601)	(0.707)	(1.853)
Geographical Region							
Northwest	-0.556	-0.744	-1.049*	-1.299+	-2.485+	-2.118	1.836
	(0.483)	(0.452)	(0.526)	(0.739)	(1.313)	(1.544)	(4.045)
Northeast	-0.889*	-1.096**	-1.364**	-1.597*	-2.892*	-2.069	1.057
	(0.435)	(0.407)	(0.473)	(0.665)	(1.181)	(1.388)	(3.637)
West	-0.550	-0.961*	-1.084*	-1.473*	-3.240**	-2.705+	-0.948
	(0.447)	(0.419)	(0.487)	(0.684)	(1.215)	(1.429)	(3.744)
Centre	-0.314	-0.695+	-1.216**	-1.522*	-4.073***	-4.028**	-2.847
	(0.416)	(0.389)	(0.452)	(0.636)	(1.129)	(1.327)	(3.477)
South	-0.169	0.109	0.042	-0.954	-3.370*	-3.705*	-0.911
	(0.542)	(0.507)	(0.590)	(0.829)	(1.472)	(1.731)	(4.535)
Constant	20.075***	21.449***	24.212***	27.848***	33.016***	37.114***	40.923***
	(0.604)	(0.565)	(0.657)	(0.923)	(1.640)	(1.928)	(5.051)
N	5,103	5,103	5,103	5,103	5,103	5,103	5,103
r2	.00796	.00616	.00992	.00953	.0113	.0113	.00504

Notes: standard errors in parenthesis + $p<0.1$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$
r2=R squared

Table B.6.9: Linear regression: *Ex-Ante* IOp in WC across different percentiles. 2006

	p10	**p25**	**p50**	**p75**	**p90**	**p95**	**p99**
Individual Characteristics							
Individual's sex	2.724***	3.395***	2.766***	3.728***	3.121**	5.130**	4.669+
	(0.547)	(0.537)	(0.587)	(0.754)	(1.131)	(1.840)	(2.772)
Individual's ethnicity	0.233	4.619**	2.922+	7.159***	5.627+	4.041	-10.165
	(1.492)	(1.462)	(1.600)	(2.054)	(3.081)	(5.014)	(7.555)
Household living standards							
Running water HH	0.759	0.686	-0.117	0.316	-2.235	-2.404	-4.729
	(0.738)	(0.724)	(0.792)	(1.017)	(1.525)	(2.482)	(3.739)
LS:Low	0.483	0.898	1.429	1.330	2.900	-0.393	10.377*
	(0.896)	(0.878)	(0.961)	(1.234)	(1.851)	(3.011)	(4.537)
LS:Med-Low	1.163	0.985	1.847*	2.203*	-0.167	-2.167	6.373
	(0.770)	(0.755)	(0.826)	(1.060)	(1.590)	(2.588)	(3.899)
LS:Med-High	-0.829	0.338	0.460	1.125	0.405	-0.836	4.954
	(0.757)	(0.742)	(0.811)	(1.042)	(1.563)	(2.543)	(3.831)
State Deprivation							
Dep:V.High	-0.918	-5.896**	-6.165**	-3.952	-4.752	-4.160	-8.142
	(2.031)	(1.990)	(2.178)	(2.796)	(4.194)	(6.824)	(10.282)
Dep:High	1.004	-0.531	-1.150	-1.417	-4.008*	-5.072	-9.722*
	(0.966)	(0.946)	(1.036)	(1.330)	(1.994)	(3.245)	(4.889)
Dep:Med	-1.183	-0.809	-1.100	-1.802	-4.744*	-6.457*	-10.138*
	(0.914)	(0.895)	(0.980)	(1.258)	(1.887)	(3.070)	(4.626)
Dep:Low	0.119	-0.309	0.454	0.501	-2.080	-2.821	-3.618
	(0.843)	(0.826)	(0.904)	(1.161)	(1.741)	(2.833)	(4.268)
Geographical Region							
Northwest	1.829	1.486	0.147	-0.753	1.696	1.853	2.972
	(1.956)	(1.917)	(2.098)	(2.694)	(4.040)	(6.574)	(9.906)
Northeast	0.550	1.052	1.521	-0.445	-0.101	5.998	1.161
	(1.814)	(1.778)	(1.946)	(2.498)	(3.747)	(6.098)	(9.187)
West	1.446	1.323	-0.288	-3.566	-2.689	-0.671	-0.539
	(1.862)	(1.824)	(1.996)	(2.563)	(3.845)	(6.256)	(9.426)
Centre	-0.158	-0.739	-1.624	-4.027+	-4.678	-5.043	-6.235
	(1.747)	(1.712)	(1.874)	(2.406)	(3.609)	(5.872)	(8.847)
South	1.165	4.402*	1.561	-0.285	-1.504	1.302	-0.574
	(2.288)	(2.242)	(2.453)	(3.150)	(4.725)	(7.688)	(11.583)
Constant	69.298***	70.019***	80.082***	85.511***	100.085***	109.488***	135.107***
	(2.360)	(2.313)	(2.531)	(3.250)	(4.875)	(7.932)	(11.951)
N	2,674	2,674	2,674	2,674	2,674	2,674	2,674
r2	.0168	.0281	.0264	.0259	.0167	.0179	.00868

Notes: standard errors in parenthesis + $p<0.1$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$
r2=R squared

Table B.6.10: Linear regression: *Ex-Ante* IOp in BMI across different percentiles. 2012

	p10	**p25**	**p50**	**p75**	**p90**	**p95**	**p99**
Individual Characteristics							
Individual's sex	0.308	0.364	-0.407	-0.006	-0.878	-1.016	-1.525
	(0.375)	(0.370)	(0.422)	(0.480)	(0.683)	(1.037)	(1.156)
Individual's ethnicity	-0.742	0.579	0.037	-0.090	0.056	-0.373	0.476
	(0.722)	(0.711)	(0.812)	(0.925)	(1.315)	(1.995)	(2.225)
Household living standards							
Running water HH	-0.288	0.230	0.789	0.125	-2.234*	-1.860	1.178
	(0.521)	(0.513)	(0.585)	(0.666)	(0.948)	(1.438)	(1.603)
LS:Low	-0.945	-0.477	-0.101	-0.835	-4.089**	-4.513*	-1.765
	(0.708)	(0.697)	(0.795)	(0.906)	(1.289)	(1.955)	(2.180)
LS:Med-Low	-0.128	0.879	1.767*	0.377	-0.489	0.545	-2.817
	(0.621)	(0.611)	(0.698)	(0.795)	(1.130)	(1.715)	(1.912)
LS:Med-High	0.459	0.506	0.381	-0.773	-1.376	-3.008+	-3.983*
	(0.565)	(0.556)	(0.635)	(0.723)	(1.029)	(1.561)	(1.740)
State Deprivation							
Dep:V.High	1.601	2.817*	1.884	-0.893	-2.859	-1.801	-0.859
	(1.298)	(1.279)	(1.459)	(1.662)	(2.364)	(3.587)	(4.000)
Dep:High	-0.035	-0.752	-1.061	-1.406	-2.823*	0.749	-1.986
	(0.705)	(0.695)	(0.793)	(0.903)	(1.284)	(1.948)	(2.172)
Dep:Med	0.292	-0.679	-0.910	-1.589*	-1.977+	1.926	-0.616
	(0.622)	(0.612)	(0.699)	(0.796)	(1.132)	(1.717)	(1.915)
Dep:Low	0.643	-0.112	-0.297	-1.343+	-0.908	-0.087	-3.267+
	(0.624)	(0.615)	(0.701)	(0.799)	(1.136)	(1.724)	(1.923)
Geographical Region							
Northwest	-1.162	-2.637*	-0.995	0.681	0.024	2.367	-2.283
	(1.270)	(1.250)	(1.427)	(1.625)	(2.312)	(3.507)	(3.911)
Northeast	-1.623	-2.995**	-2.208+	-1.360	-2.215	-2.622	-6.605+
	(1.122)	(1.105)	(1.261)	(1.437)	(2.044)	(3.100)	(3.457)
West	-0.593	-3.271**	-1.923	-0.672	-2.405	-2.464	-6.018+
	(1.172)	(1.154)	(1.317)	(1.501)	(2.134)	(3.238)	(3.611)
Centre	0.307	-1.680	-1.227	-0.726	-2.434	-1.075	-6.199+
	(1.085)	(1.068)	(1.219)	(1.388)	(1.975)	(2.996)	(3.341)
South	-0.248	-3.363*	-3.048*	-1.726	-1.602	0.143	-6.471
	(1.374)	(1.353)	(1.544)	(1.759)	(2.502)	(3.796)	(4.233)
Constant	22.124***	24.536***	27.323***	32.292***	40.195***	41.168***	50.449***
	(1.547)	(1.524)	(1.739)	(1.980)	(2.817)	(4.273)	(4.765)
N	1,045	1,045	1,045	1,045	1,045	1,045	1,045
r2	.0296	.0363	.0287	.0167	.0333	.0279	.0186

Notes: standard errors in parenthesis + $p<0.1$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$
r2=R squared

Table B.6.11: Linear regression: *Ex-Ante* IOp in WC across different percentiles. 2012

	p10	**p25**	**p50**	**p75**	**p90**	**p95**	**p99**
Individual Characteristics							
Individual's sex	2.397**	3.092**	3.464**	4.794***	6.129***	5.734*	1.421
	(0.921)	(1.004)	(1.077)	(1.137)	(1.810)	(2.632)	(4.297)
Individual's ethnicity	0.101	-1.324	1.388	2.312	1.377	-2.808	-0.197
	(1.813)	(1.975)	(2.120)	(2.238)	(3.563)	(5.180)	(8.456)
Household living standards							
Running water HH	0.858	-0.070	-0.271	2.046	1.390	-2.936	-8.257
	(1.291)	(1.407)	(1.510)	(1.594)	(2.538)	(3.689)	(6.023)
LS:Low	-1.734	-0.858	-2.848	-1.731	-7.990*	-13.527**	-3.035
	(1.743)	(1.899)	(2.038)	(2.151)	(3.425)	(4.979)	(8.129)
LS:Med-Low	-0.071	2.678	2.271	1.917	-1.800	-3.340	2.625
	(1.526)	(1.663)	(1.785)	(1.884)	(3.000)	(4.361)	(7.120)
LS:Med-High	0.503	1.241	-0.180	-3.163+	-8.138**	-6.069	2.807
	(1.380)	(1.503)	(1.614)	(1.703)	(2.712)	(3.942)	(6.436)
State Deprivation							
Dep:V.High	7.508*	9.050**	3.038	-2.777	-8.279	-12.271	-18.077
	(3.186)	(3.471)	(3.726)	(3.933)	(6.261)	(9.103)	(14.861)
Dep:High	0.642	-0.058	-1.723	-3.805+	-3.456	-3.168	-3.421
	(1.728)	(1.883)	(2.022)	(2.133)	(3.397)	(4.938)	(8.062)
Dep:Med	2.225	0.255	-0.610	-3.913*	0.252	0.615	-3.364
	(1.535)	(1.673)	(1.795)	(1.895)	(3.017)	(4.386)	(7.160)
Dep:Low	2.028	1.553	0.478	-1.837	-2.182	-2.708	-7.090
	(1.533)	(1.670)	(1.793)	(1.892)	(3.012)	(4.379)	(7.149)
Geographical Region							
Northwest	-0.911	-1.142	-0.981	-0.171	3.739	1.100	3.790
	(3.125)	(3.405)	(3.655)	(3.857)	(6.141)	(8.928)	(14.576)
Northeast	-0.491	-2.704	-2.521	-3.142	-2.960	-4.282	-1.157
	(2.785)	(3.035)	(3.258)	(3.438)	(5.474)	(7.958)	(12.992)
West	-2.318	-3.747	-2.484	-1.191	-1.433	-5.493	-0.683
	(2.900)	(3.160)	(3.392)	(3.580)	(5.700)	(8.286)	(13.528)
Centre	-0.985	-1.557	-2.913	-2.345	-3.546	-4.076	-8.842
	(2.695)	(2.937)	(3.153)	(3.327)	(5.297)	(7.701)	(12.573)
South	-3.226	-7.610*	-6.476	-4.609	-0.100	-0.755	3.700
	(3.395)	(3.699)	(3.971)	(4.191)	(6.672)	(9.700)	(15.837)
Constant	71.931***	79.792***	88.064***	96.307***	108.931***	127.678***	134.706***
	(3.844)	(4.189)	(4.497)	(4.746)	(7.556)	(10.985)	(17.934)
N	980	980	980	980	980	980	980
r2	.0231	.0304	.0351	.052	.0475	.0269	.0092

Notes: standard errors in parenthesis + p<0.1, * p<0.05, ** p<0.01, *** p<0.001
r2=R squared

Table B.6.12: Linear regression: *Ex-Ante* IOp in BMI across different percentiles. 2018

	p10	**p25**	**p50**	**p75**	**p90**	**p95**	**p99**
Individual Characteristics							
Individual's sex	0.569+	0.016	-0.302	-0.812**	-1.579**	-2.006**	-2.478*
	(0.323)	(0.269)	(0.261)	(0.314)	(0.498)	(0.648)	(1.135)
Individual's ethnicity	-0.216	-0.002	0.547	1.633**	2.334*	1.799	0.670
	(0.633)	(0.526)	(0.512)	(0.615)	(0.976)	(1.269)	(2.224)
Household living standards							
Running water HH	-0.365	-0.344	-0.533	-1.306**	-0.858	-0.446	1.860
	(0.477)	(0.397)	(0.387)	(0.464)	(0.737)	(0.958)	(1.678)
LS:Low	1.490*	0.951+	0.441	0.853	-0.218	-1.771	-2.091
	(0.638)	(0.531)	(0.517)	(0.620)	(0.985)	(1.281)	(2.244)
LS:Med-Low	0.902	1.006*	0.711	1.526**	-0.156	-0.554	-0.799
	(0.591)	(0.492)	(0.479)	(0.575)	(0.913)	(1.187)	(2.080)
LS:Med-High	0.882+	0.950*	0.821*	0.955+	1.206	0.619	3.043+
	(0.503)	(0.419)	(0.408)	(0.489)	(0.777)	(1.010)	(1.770)
State Deprivation							
Dep:V.High	2.824*	1.764+	-0.222	-0.645	-1.745	-4.116	-11.930**
	(1.258)	(1.047)	(1.018)	(1.222)	(1.941)	(2.524)	(4.422)
Dep:High	1.724*	1.169*	0.088	-1.299+	-3.377**	-6.021***	-12.944***
	(0.701)	(0.583)	(0.568)	(0.681)	(1.082)	(1.407)	(2.465)
Dep:Med	1.322*	1.234*	0.574	-0.539	-2.720**	-4.757***	-12.435***
	(0.645)	(0.537)	(0.522)	(0.627)	(0.996)	(1.295)	(2.269)
Dep:Low	0.645	0.454	-0.076	-1.777**	-3.491***	-5.267***	-10.635***
	(0.682)	(0.567)	(0.552)	(0.662)	(1.052)	(1.368)	(2.397)
Geographical Region							
Northwest	-0.569	-1.828+	-2.789**	-2.061+	-3.677*	-3.316	-3.446
	(1.126)	(0.937)	(0.912)	(1.095)	(1.738)	(2.260)	(3.961)
Northeast	-0.692	-1.420+	-2.419**	-3.171***	-5.227***	-4.387*	-3.925
	(0.984)	(0.819)	(0.797)	(0.956)	(1.518)	(1.974)	(3.459)
West	-1.272	-1.851*	-2.476**	-2.838**	-5.212**	-4.737*	-4.440
	(1.033)	(0.860)	(0.836)	(1.004)	(1.594)	(2.073)	(3.632)
Centre	-0.969	-2.183**	-2.778***	-3.535***	-6.149***	-5.331**	-4.286
	(0.943)	(0.785)	(0.764)	(0.917)	(1.456)	(1.893)	(3.317)
South	-2.596*	-1.910+	-1.796+	-3.438**	-7.121***	-6.610*	-4.450
	(1.300)	(1.082)	(1.053)	(1.263)	(2.006)	(2.609)	(4.572)
Constant	21.673***	25.412***	29.638***	34.132***	41.991***	47.350***	57.653***
	(1.263)	(1.051)	(1.023)	(1.227)	(1.949)	(2.535)	(4.442)
N	2,109	2,109	2,109	2,109	2,109	2,109	2,109
r2	.0102	.013	.0131	.0266	.028	.0239	.0249

Notes: standard errors in parenthesis + p<0.1, * p<0.05, ** p<0.01, *** p<0.001
r2=R squared

Table B.6.13: Linear regression: *Ex-Ante* IOp in WC across different percentiles. 2018

	p10	**p25**	**p50**	**p75**	**p90**	**p95**	**p99**
Individual Characteristics							
Individual's sex	2.681***	4.111***	2.892***	2.966***	5.004***	4.033*	0.839
	(0.752)	(0.722)	(0.635)	(0.746)	(1.206)	(1.762)	(2.927)
Individual's ethnicity	-1.587	2.304	2.834*	4.563**	2.948	-1.212	-15.708**
	(1.475)	(1.417)	(1.245)	(1.463)	(2.366)	(3.456)	(5.744)
Household living standards							
Running water HH	-1.845+	-0.795	-1.071	-0.507	-0.887	2.175	17.104***
	(1.114)	(1.070)	(0.940)	(1.105)	(1.787)	(2.611)	(4.338)
LS:Low	4.321**	0.667	0.186	0.988	2.733	0.194	-17.903**
	(1.487)	(1.429)	(1.255)	(1.475)	(2.386)	(3.485)	(5.791)
LS:Med-Low	3.264*	0.830	1.249	2.516+	3.981+	0.881	-16.952**
	(1.380)	(1.326)	(1.165)	(1.369)	(2.214)	(3.234)	(5.374)
LS:Med-High	2.804*	1.127	0.923	2.362*	2.475	0.345	-5.332
	(1.172)	(1.126)	(0.990)	(1.163)	(1.881)	(2.747)	(4.565)
State Deprivation							
Dep:V.High	-3.575	-2.929	-3.266	0.061	1.421	-10.564	-36.150**
	(2.920)	(2.805)	(2.464)	(2.895)	(4.684)	(6.841)	(11.368)
Dep:High	1.908	-0.525	-1.928	-2.137	-2.858	-12.113**	-22.924***
	(1.625)	(1.561)	(1.372)	(1.612)	(2.607)	(3.809)	(6.329)
Dep:Med	3.287*	2.207	0.794	0.843	-2.411	-8.838*	-23.389***
	(1.499)	(1.440)	(1.266)	(1.487)	(2.405)	(3.513)	(5.838)
Dep:Low	2.555	2.217	-0.964	-0.393	-2.284	-11.496**	-23.844***
	(1.576)	(1.514)	(1.330)	(1.563)	(2.528)	(3.693)	(6.137)
Geographical Region							
Northwest	-0.820	-3.989	-1.226	-1.775	-2.696	-1.255	-4.470
	(2.643)	(2.539)	(2.231)	(2.621)	(4.240)	(6.193)	(10.291)
Northeast	0.650	-0.730	-1.238	-1.891	-4.778	-5.172	-4.586
	(2.297)	(2.207)	(1.939)	(2.278)	(3.685)	(5.383)	(8.944)
West	-0.517	-2.864	-1.389	0.162	-3.878	-1.629	4.428
	(2.412)	(2.317)	(2.036)	(2.392)	(3.869)	(5.651)	(9.391)
Centre	-0.077	-2.778	-2.669	-2.135	-4.267	-2.804	-0.805
	(2.200)	(2.114)	(1.857)	(2.182)	(3.530)	(5.156)	(8.568)
South	4.243	1.735	-0.127	-1.773	-6.286	-2.019	11.443
	(3.025)	(2.906)	(2.554)	(3.000)	(4.853)	(7.089)	(11.780)
Constant	75.726***	82.491***	91.790***	96.216***	109.215***	126.787***	165.706***
	(2.944)	(2.828)	(2.485)	(2.919)	(4.723)	(6.898)	(11.463)
N	2,055	2,055	2,055	2,055	2,055	2,055	2,055
r2	.0161	.0279	.0243	.0276	.0138	.00985	.0239

Notes: standard errors in parenthesis + p<0.1, * p<0.05, ** p<0.01, *** p<0.001
r2=R squared

B.7 Linear regression models. *Ex-post* approach

Table B.7.1: Linear regression results: *Ex-post* approach. Stage I. 2006

	DP1	DP2	DP3	FS	PA	Alc.	Tob.
Individual Characteristics							
Individual's sex	2.145***	-1.189***	2.735***	0.013	0.450**	0.216***	0.235***
	(0.289)	(0.260)	(0.306)	(0.011)	(0.143)	(0.013)	(0.012)
Individual's ethnicity	2.850***	-0.348	0.623	0.029	-0.437	0.085*	0.012
	(0.652)	(0.587)	(0.689)	(0.026)	(0.416)	(0.036)	(0.034)
Household living standards							
Running water HH	-0.516	-0.433	0.736+	-0.004	-0.034	-0.060***	-0.013
	(0.363)	(0.327)	(0.384)	(0.014)	(0.202)	(0.018)	(0.016)
LS:Low	-4.420***	-1.065*	3.897***	-0.023	0.858***	-0.111***	-0.102***
	(0.485)	(0.437)	(0.513)	(0.019)	(0.241)	(0.021)	(0.020)
LS:Med-Low	-3.132***	-1.322**	3.445***	-0.006	0.553**	-0.112***	-0.000
	(0.458)	(0.412)	(0.484)	(0.018)	(0.206)	(0.018)	(0.017)
LS:Med-High	-2.172***	-0.876*	1.696***	-0.019	0.407*	-0.059***	0.009
	(0.458)	(0.412)	(0.484)	(0.018)	(0.199)	(0.017)	(0.016)
State Deprivation							
Dep:V.High	-1.024	2.493**	3.164**	0.064	0.401	-0.021	-0.026
	(1.032)	(0.929)	(1.090)	(0.040)	(0.529)	(0.046)	(0.043)
Dep:High	-1.662**	0.855+	2.344***	-0.014	0.402	-0.005	-0.041+
	(0.555)	(0.500)	(0.587)	(0.022)	(0.265)	(0.023)	(0.022)
Dep:Med	-1.941***	1.448**	2.300***	0.005	-0.343	-0.058**	-0.096***
	(0.534)	(0.481)	(0.565)	(0.021)	(0.249)	(0.022)	(0.020)
Dep:Low	-0.148	1.071*	0.658	0.009	-0.384+	-0.004	-0.016
	(0.504)	(0.454)	(0.533)	(0.020)	(0.233)	(0.020)	(0.019)
Geographical Region							
Northwest	-2.310*	0.839	-0.557	-0.039	0.654	0.068	0.023
	(1.018)	(0.916)	(1.076)	(0.040)	(0.507)	(0.044)	(0.041)
Northeast	-1.959*	-0.210	1.899*	0.005	0.394	0.014	0.043
	(0.893)	(0.804)	(0.943)	(0.035)	(0.455)	(0.040)	(0.037)
West	-2.716**	-0.036	0.760	-0.007	0.568	0.069+	0.065+
	(0.918)	(0.827)	(0.971)	(0.036)	(0.469)	(0.041)	(0.038)
Centre	-3.106***	1.710*	-0.062	0.005	0.086	-0.013	0.079*
	(0.845)	(0.761)	(0.894)	(0.033)	(0.435)	(0.038)	(0.035)
South	-3.877***	-1.433	-0.402	-0.042	-0.675	-0.030	-0.016
	(1.110)	(0.999)	(1.173)	(0.044)	(0.569)	(0.050)	(0.046)
Constant	26.433***	21.470***	25.100***	0.049	3.915***	0.214***	0.136**
	(1.149)	(1.035)	(1.215)	(0.045)	(0.632)	(0.055)	(0.051)
N	2,040	2,040	2,040	2,040	5,126	5,126	5,125
r2	.165	.0421	.114	.00838	.00931	.0884	.102

Notes: standard errors in parenthesis + p<0.1, * p<0.05, ** p<0.01, *** p<0.001
DP1= low-nutritious and high-energy food. DP2=high-nutritious food. DP3=legumes and maize-based products. FS=food supplements. PA=physical activity. Alc=Alcohol consumption.Tob.=Smoking
Phase I means that effort variables are regressed against circumstances.

Table B.7.2: Logit regression results: *Ex-post* approach. Stage II. 2006

	DBM	**HB**	**BMI**	**WC**
Individual Characteristics				
Individual's sex	-1.912***	-1.616***	-0.003	-1.224***
	(0.301)	(0.200)	(0.094)	(0.144)
Individual's ethnicity	0.763	0.991*	0.724**	0.632*
	(0.621)	(0.484)	(0.242)	(0.297)
Household living standards				
Running water HH	-0.183	-0.044	-0.022	-0.236
	(0.281)	(0.198)	(0.119)	(0.163)
LS:Low	0.855*	0.595*	-0.307*	0.056
	(0.393)	(0.282)	(0.157)	(0.229)
LS:Med-Low	0.659+	0.508+	-0.065	-0.017
	(0.389)	(0.271)	(0.146)	(0.216)
LS:Med-High	-0.010	0.219	-0.340*	-0.307
	(0.410)	(0.282)	(0.148)	(0.223)
State Deprivation				
Dep:V.High	-2.669***	-0.688	-0.960**	-1.505**
	(0.803)	(0.542)	(0.330)	(0.475)
Dep:High	-0.639	-0.103	-0.342+	-0.309
	(0.420)	(0.310)	(0.180)	(0.257)
Dep:Med	-0.655+	-0.378	-0.232	-0.135
	(0.390)	(0.293)	(0.172)	(0.242)
Dep:Low	-0.294	0.035	0.006	0.032
	(0.380)	(0.279)	(0.161)	(0.234)
Geographical Region				
Northwest	0.011	0.298	-0.776*	0.275
	(0.781)	(0.549)	(0.332)	(0.485)
Northeast	-0.111	0.199	-0.508+	-0.052
	(0.707)	(0.484)	(0.288)	(0.441)
West	-0.362	-0.186	-0.381	-0.059
	(0.733)	(0.504)	(0.296)	(0.452)
Centre	-0.576	-0.395	-0.232	-0.078
	(0.678)	(0.466)	(0.272)	(0.422)
South	0.554	0.165	0.250	0.996+
	(0.787)	(0.572)	(0.353)	(0.538)
Dietary patterns and physical activity				
DP1_hat	-0.001	-0.008	-0.007	0.011
	(0.019)	(0.014)	(0.007)	(0.010)
DP2_hat	-0.019	-0.029+	0.017*	0.020+
	(0.021)	(0.015)	(0.008)	(0.011)
DP3_hat	0.027+	0.021+	-0.009	0.005
	(0.015)	(0.011)	(0.007)	(0.010)
FS_hat	0.369	0.380	-0.444*	-0.145
	(0.404)	(0.305)	(0.193)	(0.264)
PA_hat	0.071***	0.039**	0.032***	0.007
	(0.019)	(0.015)	(0.009)	(0.012)
Risky health behaviours				
Alc_hat	0.184	-0.190	0.159	-0.272
	(0.294)	(0.220)	(0.110)	(0.168)
Tob_hat	0.244	0.055	0.264*	0.792***
	(0.329)	(0.240)	(0.117)	(0.183)
Constant	-2.257*	-2.821***	-0.444	0.114
	(0.933)	(0.708)	(0.386)	(0.541)
N	1,223	1,938	2,034	993
r2_p	.158	.102	.0311	.082

Notes: standard errors in parenthesis + p<0.1, * p<0.05, ** p<0.01, *** p<0.001
DP1_hat= low-nutritious and high-energy food. DP2_hat=high-nutritious food.
DP3_hat=legumes and maize-based products. FS_hat=food supplements.
PA_hat=physical activity. Alc_hat=Alcohol consumption. Tob.=Smoking
Phase II outcomes are regressed against circumstances and true levels of effort.
r2_p=pseudo R squared. HB=1 if anaemia. BMI=1 if excess weight
WC=1 if excess adiposity. DBM=1 if excess weight or adiposity and anaemia

Table B.7.3: Linear regression results: *Ex-post* approach. Stage I. 2018

	DP1	**DP2**	**DP3**	**FS**	**PA**	**Alc.**	**Tob.**
Individual Characteristics							
Individual's sex	2.543***	-1.185***	2.895***	-0.066***	1.577***	0.266***	0.232***
	(0.256)	(0.305)	(0.322)	(0.015)	(0.128)	(0.016)	(0.016)
Individual's ethnicity	2.363***	0.267	0.214	0.009	-0.178	0.001	0.073*
	(0.460)	(0.548)	(0.580)	(0.026)	(0.237)	(0.029)	(0.030)
Household living standards							
Running water HH	0.813*	0.206	-1.037*	0.021	-0.019	0.031	0.049*
	(0.372)	(0.443)	(0.468)	(0.021)	(0.186)	(0.023)	(0.024)
LS:Low	2.876***	2.024***	-4.640***	0.024	-1.068***	0.047	0.021
	(0.497)	(0.593)	(0.626)	(0.029)	(0.244)	(0.030)	(0.031)
LS:Med-Low	1.928***	0.773	-1.628**	0.024	-0.635**	0.017	-0.007
	(0.455)	(0.543)	(0.574)	(0.026)	(0.227)	(0.028)	(0.029)
LS:Med-High	1.005*	0.308	-0.499	0.017	-0.259	0.034	0.006
	(0.391)	(0.467)	(0.493)	(0.023)	(0.196)	(0.024)	(0.025)
Geographical Region							
Northwest	-1.784***	1.363*	-1.616*	0.009	0.669*	-0.008	0.086**
	(0.526)	(0.628)	(0.663)	(0.030)	(0.261)	(0.032)	(0.033)
Northeast	-0.844+	0.160	0.547	0.025	-0.180	-0.036	0.048+
	(0.459)	(0.548)	(0.579)	(0.026)	(0.221)	(0.027)	(0.028)
West	-1.310**	1.600**	1.164*	0.052*	0.517*	-0.025	0.022
	(0.432)	(0.515)	(0.544)	(0.025)	(0.212)	(0.026)	(0.027)
Centre	-1.336**	0.976+	0.748	0.025	0.550*	-0.080**	0.047+
	(0.456)	(0.543)	(0.574)	(0.026)	(0.224)	(0.028)	(0.028)
South	-0.243	1.811**	4.135***	0.020	0.296	-0.083**	0.017
	(0.471)	(0.562)	(0.593)	(0.027)	(0.236)	(0.029)	(0.030)
Constant	13.861***	16.040***	38.223***	0.063*	2.391***	0.073*	-0.042
	(0.534)	(0.637)	(0.674)	(0.031)	(0.268)	(0.033)	(0.034)
N	1,651	1,651	1,651	1,651	2,197	2,150	2,192
r2	.136	.0348	.187	.0213	.0901	.131	.101

Notes: standard errors in parenthesis + p<0.1, * p<0.05, ** p<0.01, *** p<0.001
DP1= low-nutritious and high-energy food. DP2=high-nutritious food. DP3=legumes and maize-based products. FS=food supplements. PA=physical activity. Alc=Alcohol consumption.Tob.=Smoking
Phase I means that effort variables are regressed against circumstances.

Table B.7.4: Logit regression results: *Ex-post* approach. Stage II. 2018

	DBM	**HB**	**BMI**	**WC**
Individual Characteristics				
Individual's sex	-1.816***	-1.550***	0.109	-0.830***
	(0.197)	(0.146)	(0.118)	(0.129)
Individual's ethnicity	-0.215	-0.270	0.096	0.098
	(0.340)	(0.233)	(0.208)	(0.207)
Household living standards				
Running water HH	0.006	0.029	-0.051	-0.281
	(0.270)	(0.193)	(0.172)	(0.183)
LS:Low	0.179	-0.342	0.421+	0.843***
	(0.380)	(0.275)	(0.232)	(0.250)
LS:Med-Low	0.590+	0.229	0.527*	0.846***
	(0.331)	(0.236)	(0.214)	(0.230)
LS:Med-High	0.408	-0.011	0.395*	0.608**
	(0.292)	(0.206)	(0.183)	(0.195)
Geographical Region				
Northwest	-0.356	0.336	-1.123*	-0.371
	(0.666)	(0.452)	(0.489)	(0.439)
Northeast	-0.396	0.046	-0.876+	-0.073
	(0.631)	(0.416)	(0.461)	(0.400)
West	-1.086+	-0.438	-1.115*	-0.448
	(0.654)	(0.440)	(0.467)	(0.406)
Centre	-0.543	0.005	-1.027*	-0.378
	(0.608)	(0.404)	(0.450)	(0.381)
South	-0.298	-0.076	-0.555	-0.446
	(0.628)	(0.425)	(0.466)	(0.395)
Dietary patterns and physical activity				
DP1_hat	-0.012	-0.025+	0.004	-0.010
	(0.020)	(0.015)	(0.011)	(0.011)
DP2_hat	-0.006	-0.008	-0.009	-0.005
	(0.014)	(0.011)	(0.009)	(0.010)
DP3_hat	-0.063***	-0.041***	-0.033***	-0.019*
	(0.016)	(0.012)	(0.009)	(0.009)
FS	-1.152***	0.067	-0.827***	-0.928***
	(0.322)	(0.213)	(0.196)	(0.215)
PA	-0.056+	-0.006	-0.017	-0.054**
	(0.030)	(0.023)	(0.018)	(0.018)
Risky health behaviours				
Alc	-0.290	-0.081	-0.356*	-0.150
	(0.291)	(0.222)	(0.160)	(0.160)
Tob	-0.237	-0.351+	-0.046	-0.367*
	(0.268)	(0.207)	(0.154)	(0.151)
Constant	0.814	-0.503	1.622***	1.587***
	(0.651)	(0.438)	(0.476)	(0.415)
N	631	1,571	1,591	1,508
r2_p	.185	.106	.0337	.0728

Notes: standard errors in parenthesis + p<0.1, * p<0.05, ** p<0.01, *** p<0.001
DP1_hat= low-nutritious and high-energy food. DP2_hat=high-nutritious food.
DP3_hat=legumes and maize-based products. FS_hat=food supplements.
PA_hat=physical activity. Alc_hat=Alcohol consumption. Tob.=Smoking
Phase II outcomes are regressed against circumstances and true levels of effort.
r2_p=pseudo R squared. HB=1 if anaemia. BMI=1 if excess weight
WC=1 if excess adiposity. DBM=1 if excess weight or adiposity and anaemia

B.8 Additional Analyses

B.9 *Ex-ante* and *ex-post* IOp by sex

B.9.1 *Ex-ante*

Table B.9.1: *Ex-Ante* IOp in outcomes defined according to clinical thresholds for women

Survey year	**1988**			**1999**		**2006**		**2012**		**2016**		**2018**	
Expected age cohort	**0-5 yo**			**11-16 yo**		**18-23 yo**		**24-29 yo**		**28-33 yo**		**30-35 yo**	
Outcomes	**DI**	**N**	**Outcomes**	**DI**	**N**	**DI**	**N**	**DI**	**N**	**DI**	**N**	**DI**	**N**
Stunting	0.300***	3,026											
Wasting	0.226***	3,086											
Uweight	0.257***	3,152											
			Anaem.	0.195***	749	0.089***	2,577	0.172***	642	-	-	0.080***	1,112
Oweight	0.196***	3,098	EW (BMI)	0.220***	1,874	0.058***	2,673	0.074*	647	-	-	0.017	1,105
			EA (WC)	0.264***	1,615	0.059*	1,759	0.053	574	-	-	0.019	1,044
DBMc	0.355***	2,914	DBMa	0.426***	576	0.120***	1,595	0.212***	282	-	-	0.053+	485

Notes:s + p<0.1, * p<0.05, ** p<0.01, *** p<0.001. Stunting=Height-for-Age below -2 Z-scores; Wasting=Weight-for-Height below -2 Z-scores. Oweight=Body mass index above 2 Z-scores. Anaem.=Anaemia (HB=Haemoglobin <13 g/dl); EW=Excess weight (BMI=Body mass index > 25kg/m2) EA=Excess adiposity (WC=Waist circumference > 80 cm.) DBMc defined as HAZ or WHZ below -2 Z-scores and BMI above +2 Z-scores. DBMa in adults (BMI > 25kg/m2 or WC>80 cm and HB<13 g/dl). N= observations. yo=years old. DI=Dissimilarity Index

Table B.9.2: *Ex-Ante* IOp in outcomes defined according to clinical thresholds for men

Survey year	**1988**			**1999**		**2006**		**2012**		**2016**		**2018**	
Expected age cohort	**0-5 yo**			**11-16 yo**		**18-23 yo**		**24-29 yo**		**28-33 yo**		**30-35 yo**	
Outcomes	**DI**	**N**	**Outcomes**	**DI**	**N**	**DI**	**N**	**DI**	**N**	**DI**	**N**	**DI**	**N**
Stunting	0.272***	2,737											
Wasting	0.165***	2,753											
Uweight	0.274***	2,830											
			Anaem.	-	-	0.066**	2,428	0.081*	412	-	-	0.055	1,004
Oweight	0.128***	2,765	EW (BMI)	-	-	0.156***	2,320	-	-	-	-	0.264***	951
			EA (WC)	-	-	0.105**	939	0.110**	410	-	-	0.075*	997
DBMc	0.248***	2,566	DBMa	-	-	0.223***	1,476	-	-	-	-	0.300***	356

Notes:s + p<0.1, * p<0.05, ** p<0.01, *** p<0.001. Stunting=Height-for-Age below -2 Z-scores; Wasting=Weight-for-Height below -2 Z-scores.
Oweight=Body mass index above 2 Z-scores. Anaem.=Anaemia (HB=Haemoglobin <13 g/dl); EW=Excess weight (BMI=Body mass index > 25kg/m2)
EA=Excess adiposity (WC=Waist circumference > 80 cm.) DBMc defined as HAZ or WHZ below -2 Z-scores and BMI above +2 Z-scores.
DBMa in adults (BMI > 25kg/m2 or WC>80 cm and HB<13 g/dl). N= observations. yo=years old. DI=Dissimilarity Index
(-) Unable to estimate IOp for 2012 and 2016 due to small sample size

Figure B.9.1: Beyond the mean: *Ex-ante* IOp in BMI for women

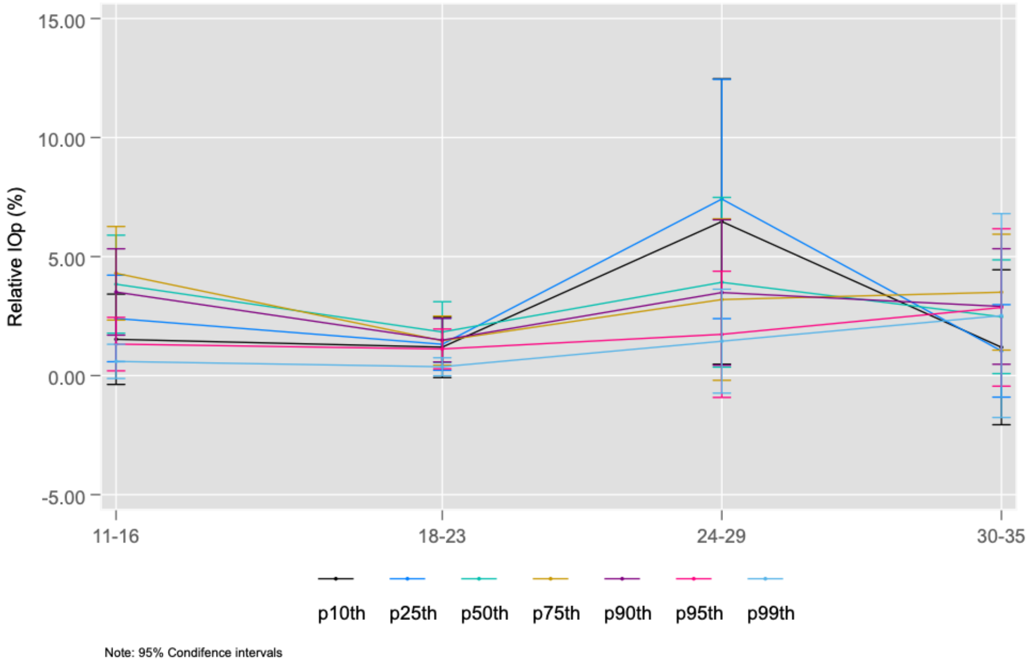


Figure B.9.2: Beyond the mean: *Ex-ante* IOp in BMI for men

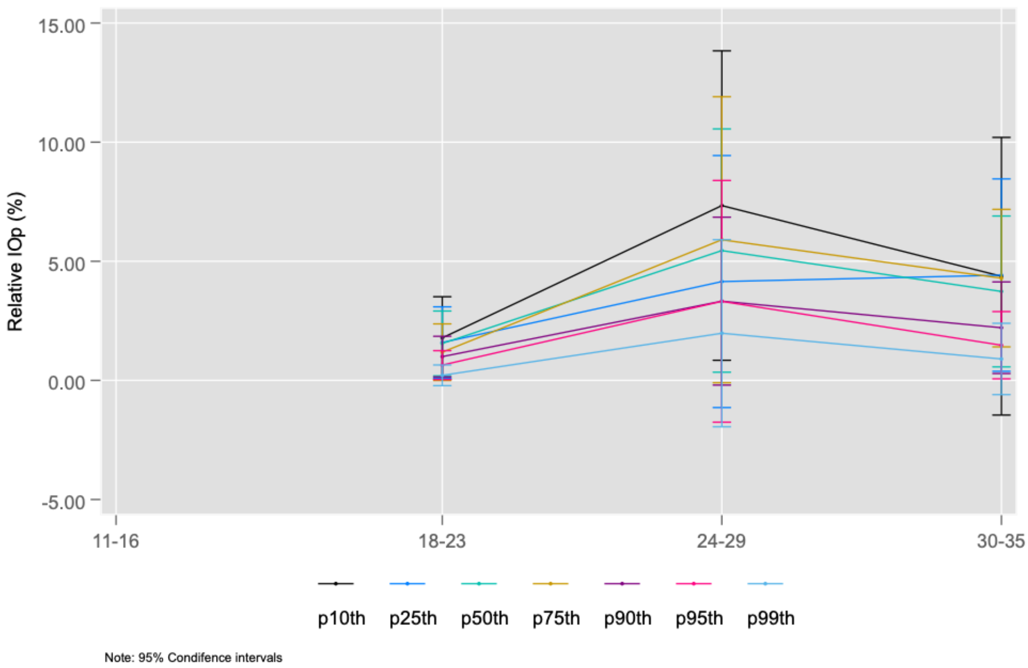


Figure B.9.3: Beyond the mean: *Ex-ante* IOp in WC for women

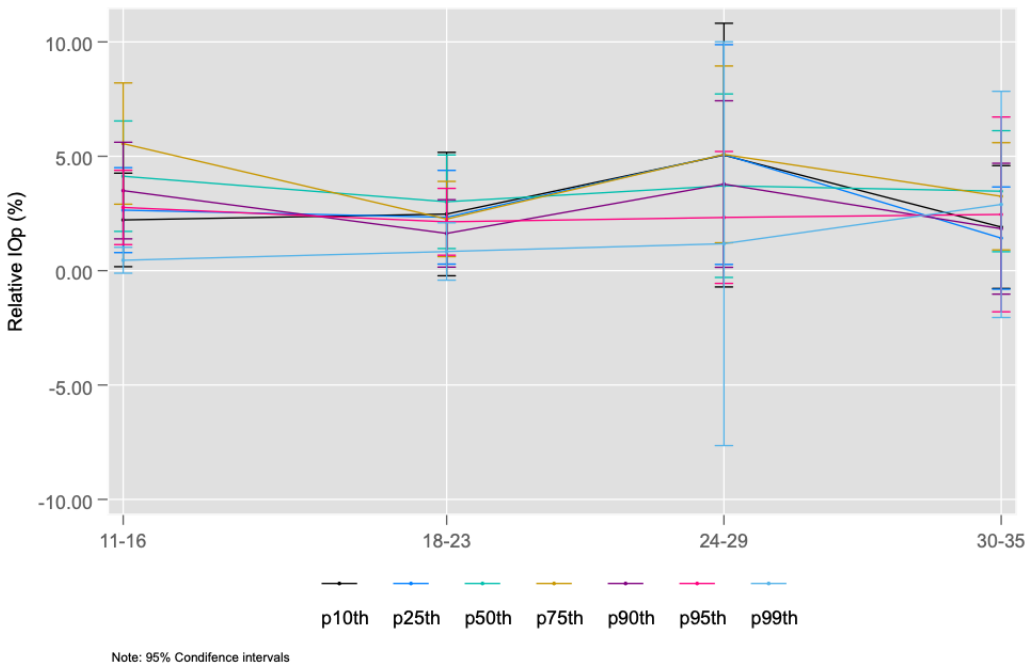
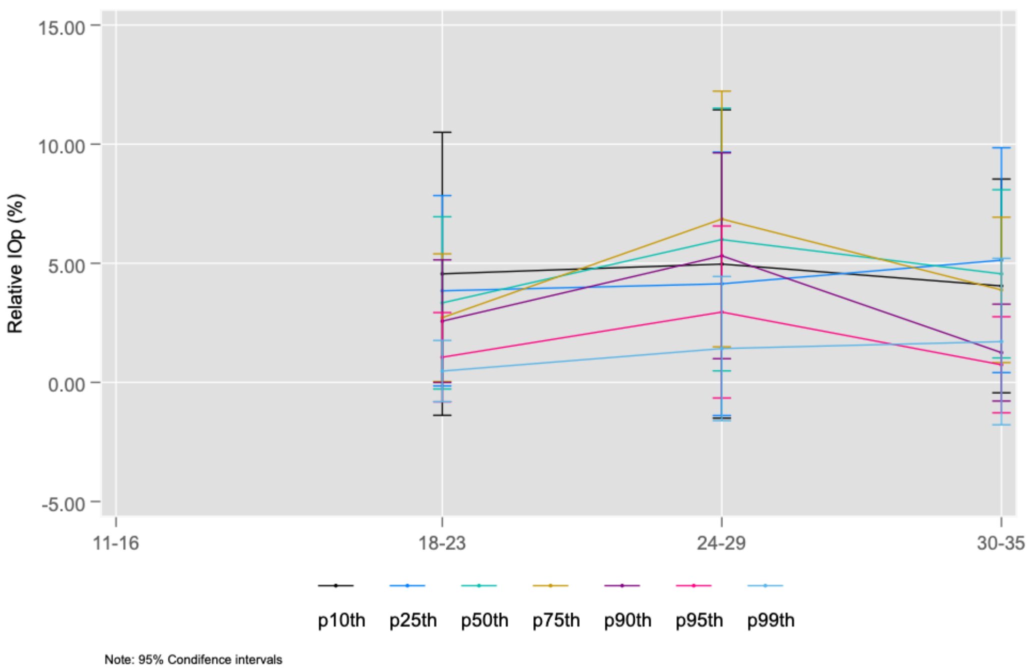


Figure B.9.4: Beyond the mean: *Ex-ante* IOp in WC for men



B.9.2 *Ex-post* IOp

Table B.9.3: *Ex-post* IOp in outcomes defined according to clinical thresholds for women

Survey year		**2006**		**2012**		**2016**		**2018**	
Expected age cohort		18-23		24-29		28-33		30-35	
Outcome		**Abs.**	**%**	**Abs.**	**%**	**Abs.**	**%**	**Abs.**	**%**
Anaem.	Circum.	0.0919	57.40	-	-	-	-	0.0653	52.38
	Efforts	0.0682	42.60	-	-	-	-	0.0594	47.62
	N		1148	-	-	-	-		951
EW (BMI)	Circum.	0.0746	70.82	-	-	-	-	0.0204	33.75
	Efforts	0.0308	29.18	-	-	-	-	0.0401	66.25
	N		1195	-	-	-	-		943
EA (WC)	Circum.	0.0521	58.00	-	-	-	-	0.0168	37.61
	Efforts	0.0377	42.00	-	-	-	-	0.0278	62.39
	N		818	-	-	-	-		883
DBM	Circum.	0.1516	51.83	-	-	-	-	0.0464	35.97
	Efforts	0.1408	48.17	-	-	-	-	0.0827	64.03
	N		715	-	-	-	-		425

Notes: N= observations. Unable to estimate IOp for 2012 and 2016 due to small sample size

Circum.=Total contribution of circumstances. Efforts=Direct contribution of efforts.

Anaem.=Anaemia (HB=Haemoglobin <13 g/dl); EW=Excess weight (BMI=Body mass index > 25kg/m2)

EA=Excess adiposity (WC=Waist circumference > 80 cm)

DBM in adults (BMI > 25kg/m2 or WC>80 cm and HB<13 g/dl)

Table B.9.4: *Ex-post* IOp in outcomes defined according to clinical thresholds for men

Survey year		**2006**		**2012**		**2016**		**2018**	
Expected age cohort		18-23		24-29		28-33		30-35	
Outcome		**Abs.**	%	**Abs.**	%	**Abs.**	%	**Abs.**	%
Anaem.	Circum.	0.3475	74.36	-	-	-	-	0.1470	53.96
	Efforts	0.1198	25.64	-	-	-	-	0.1254	46.04
	N		800	-	-	-	-		634
EW (BMI)	Circum.	0.0711	51.60	-	-	-	-	0.0471	67.90
	Efforts	0.0667	48.40	-	-	-	-	0.0223	32.10
	N		850	-	-	-	-		663
EA (WC)	Circum.	0.1157	55.37	-	-	-	-	0.0633	67.61
	Efforts	0.0933	44.63	-	-	-	-	0.0303	32.39
	N		303	-	-	-	-		657
DBM	Circum.	0.2363	74.28	-	-	-	-	0.2014	57.83
	Efforts	0.0818	25.72	-	-	-	-	0.1468	42.17
	N		309	-	-	-	-		238

Notes: N= observations. Unable to estimate IOp for 2012 and 2016 due to small sample size

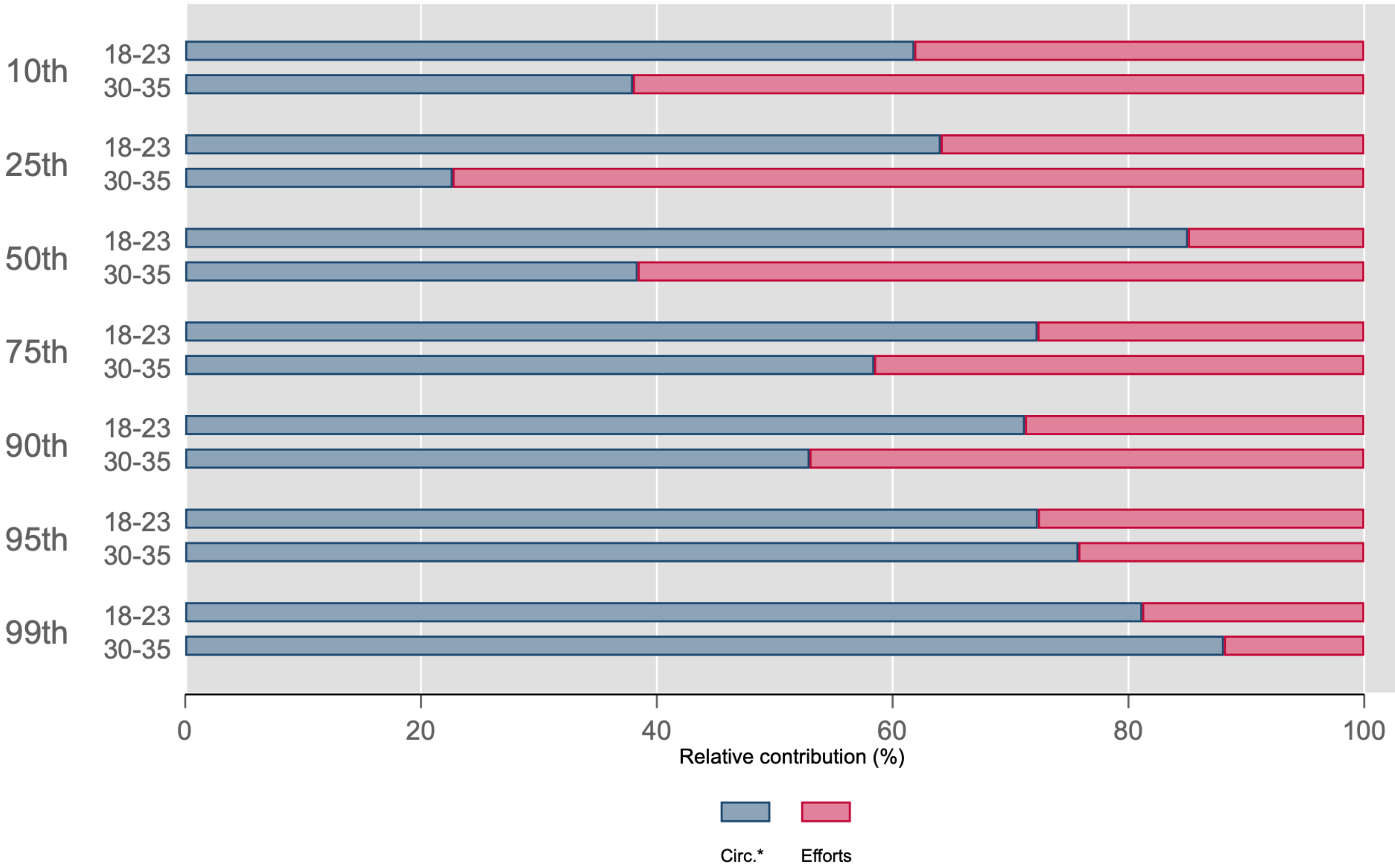
Circum.=Total contribution of circumstances. Efforts=Direct contribution of efforts.

Anaem.=Anaemia (HB=Haemoglobin <13 g/dl); EW=Excess weight (BMI=Body mass index > 25kg/m2)

EA=Excess adiposity (WC=Waist circumference > 80 cm)

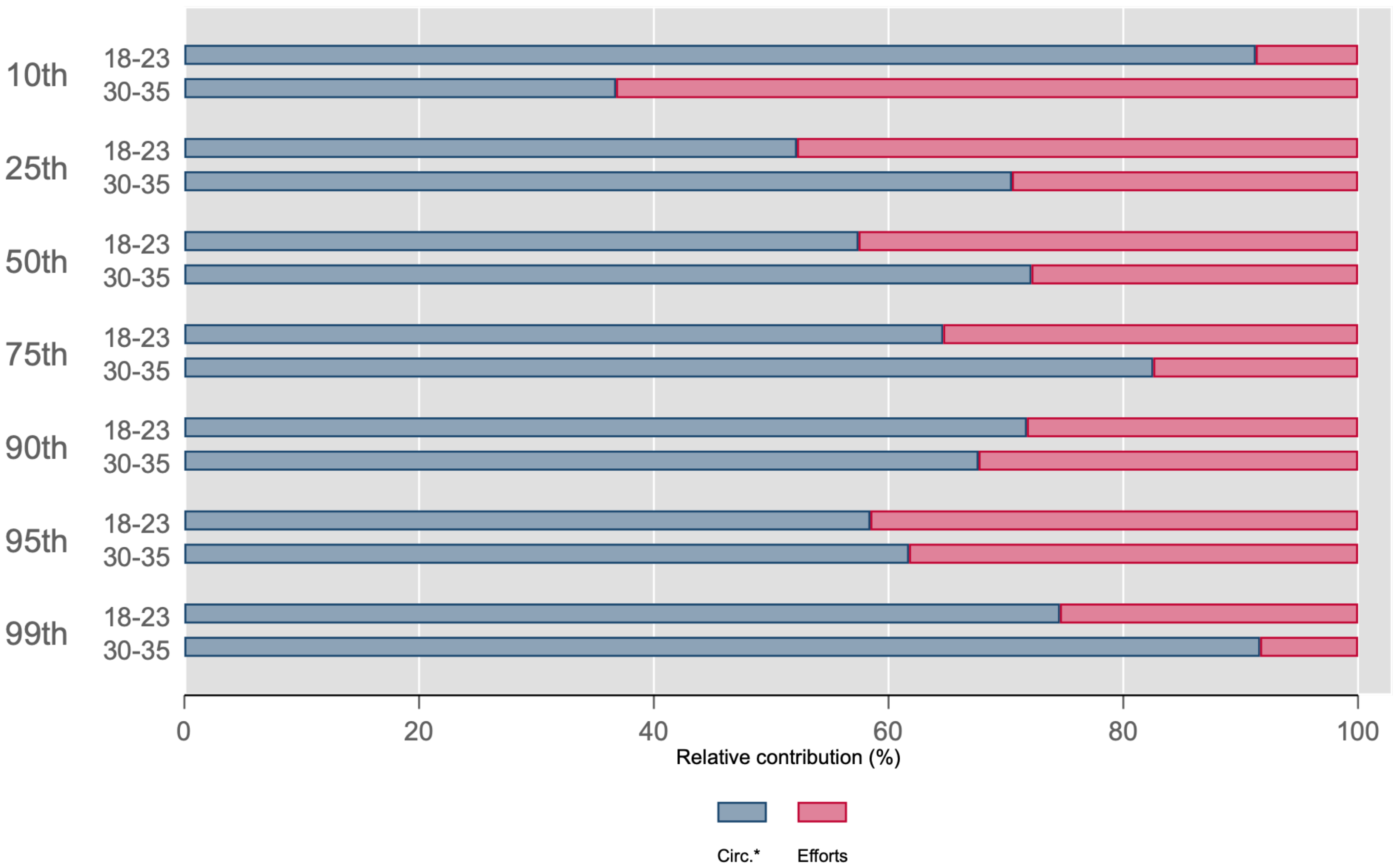
DBM in adults (BMI > 25kg/m2 or WC>80 cm and HB<13 g/dl)

Figure B.9.5: Beyond the mean: *Ex-post* IOp in BMI for women



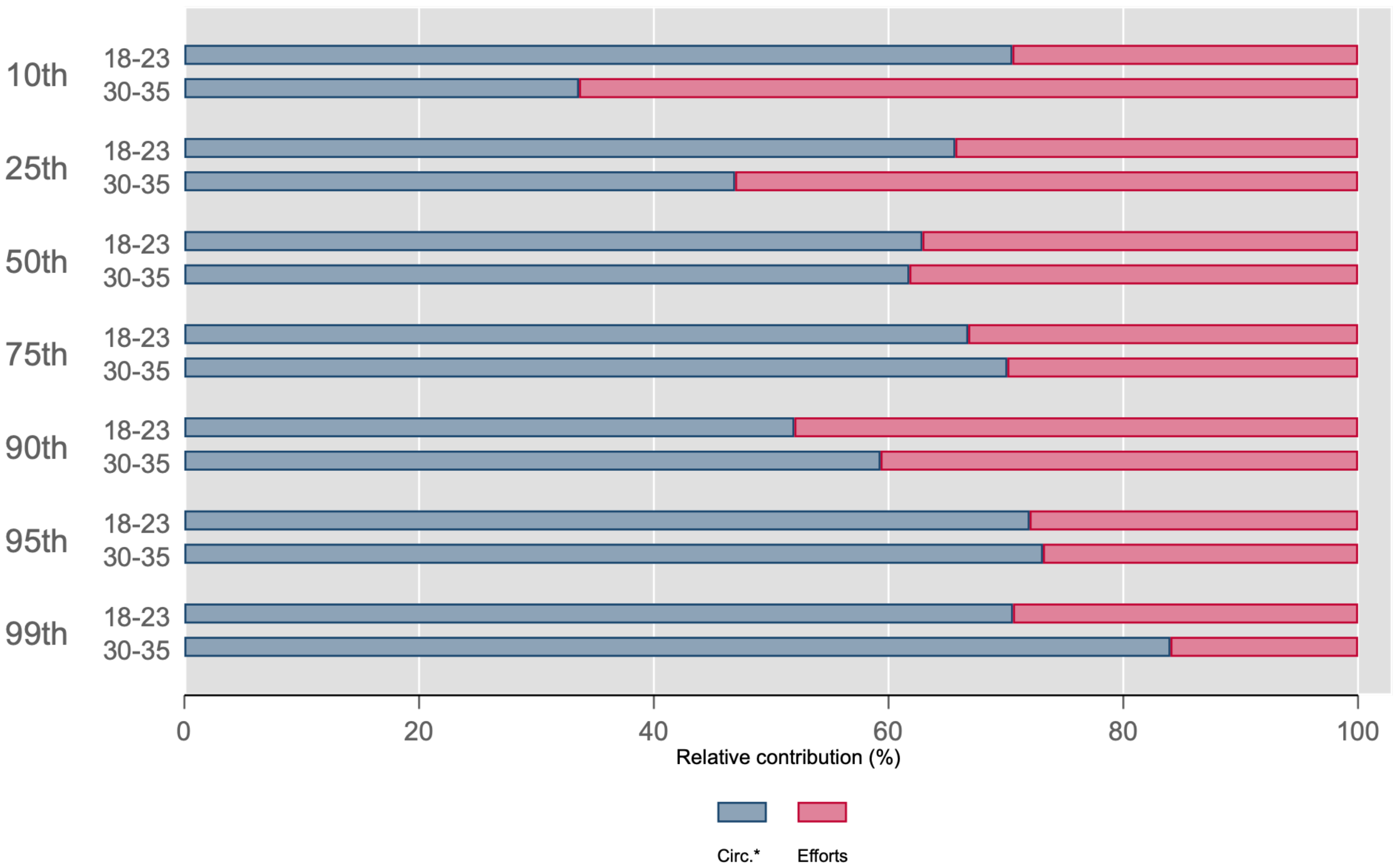
Notes: *Circ. means Circumstances. *th* indicates percentile

Figure B.9.6: Beyond the mean: *Ex-post* IOp in BMI for men



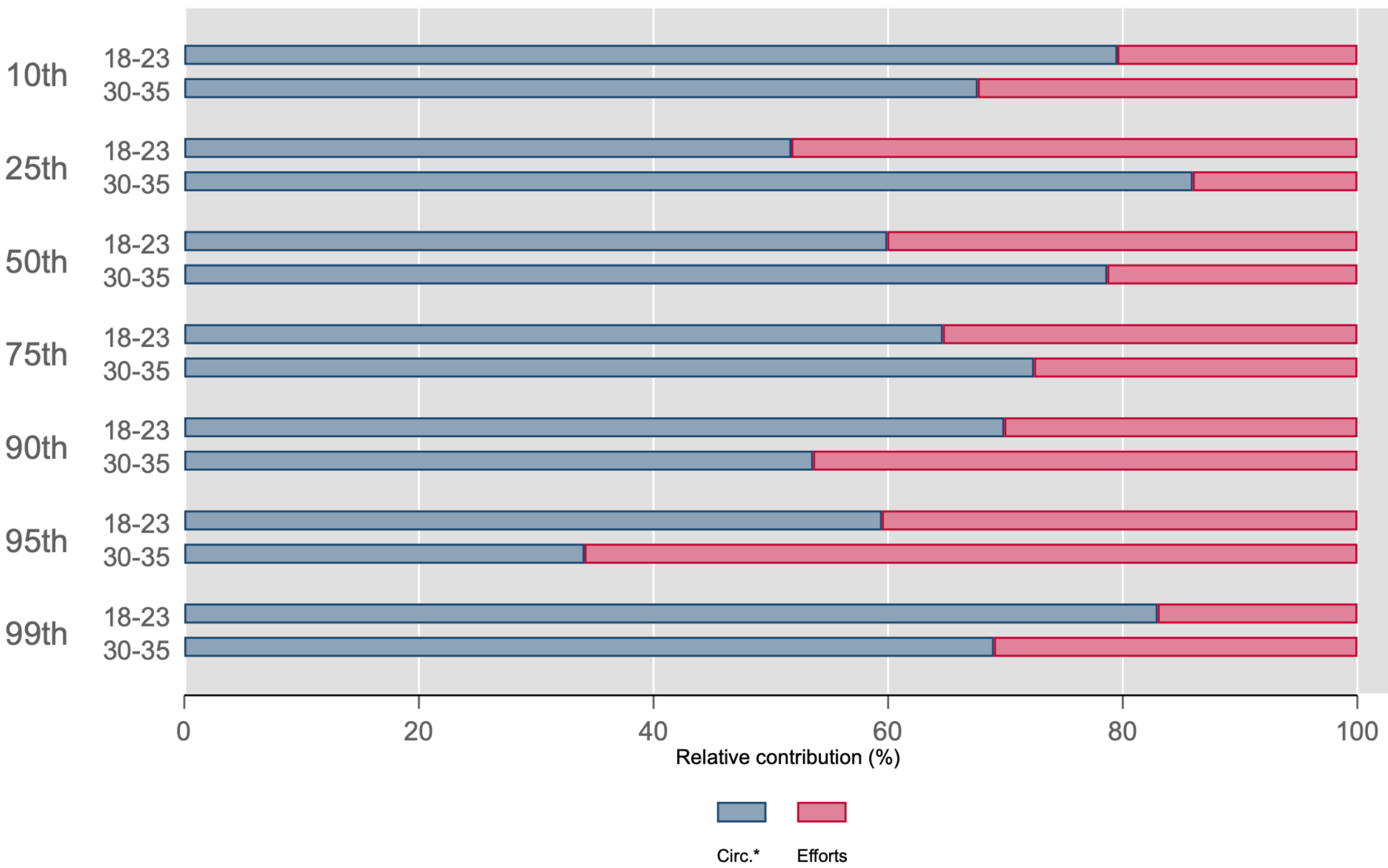
Notes: *Circ. means Circumstances. *th* indicates percentile

Figure B.9.7: Beyond the mean: *Ex-post* IOp in WC for women



Notes: *Circ. means other circumstances excluding sex. *th* indicates percentile

Figure B.9.8: Beyond the mean: *Ex-post* IOp in WC for men

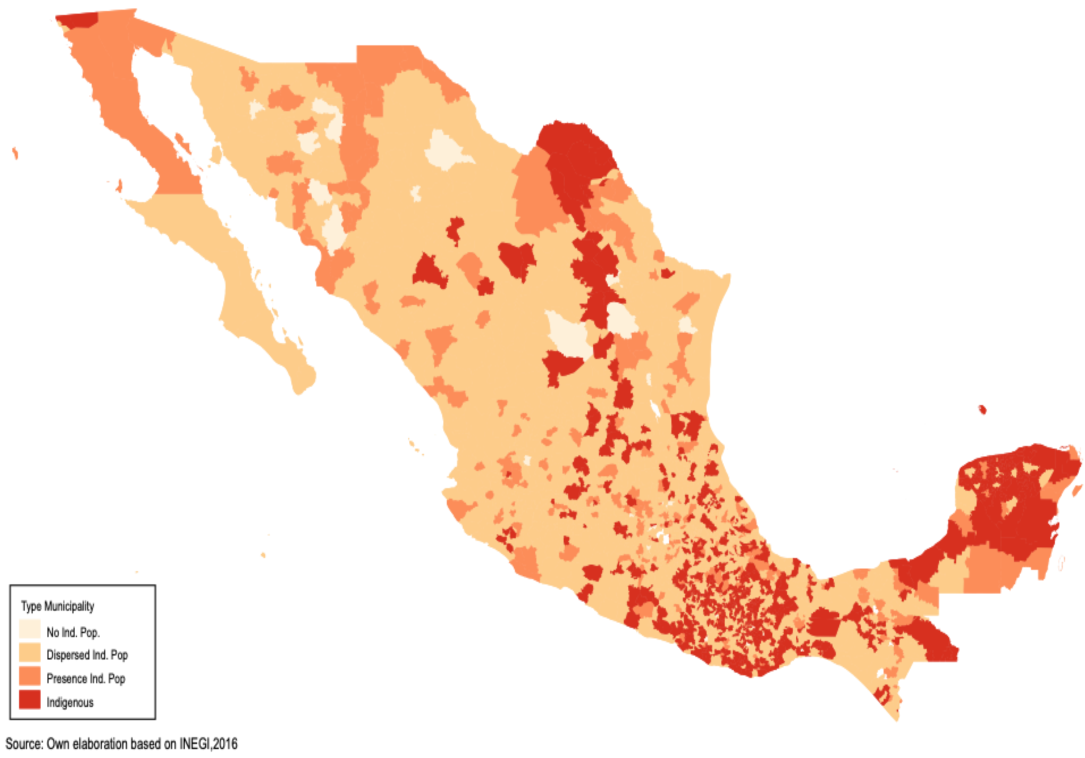


Notes: *Circ. means other circumstances excluding sex. *th* indicates percentile

Appendix C

Chapter 3

Figure C.0.1: Distribution of indigenous people across Mexico



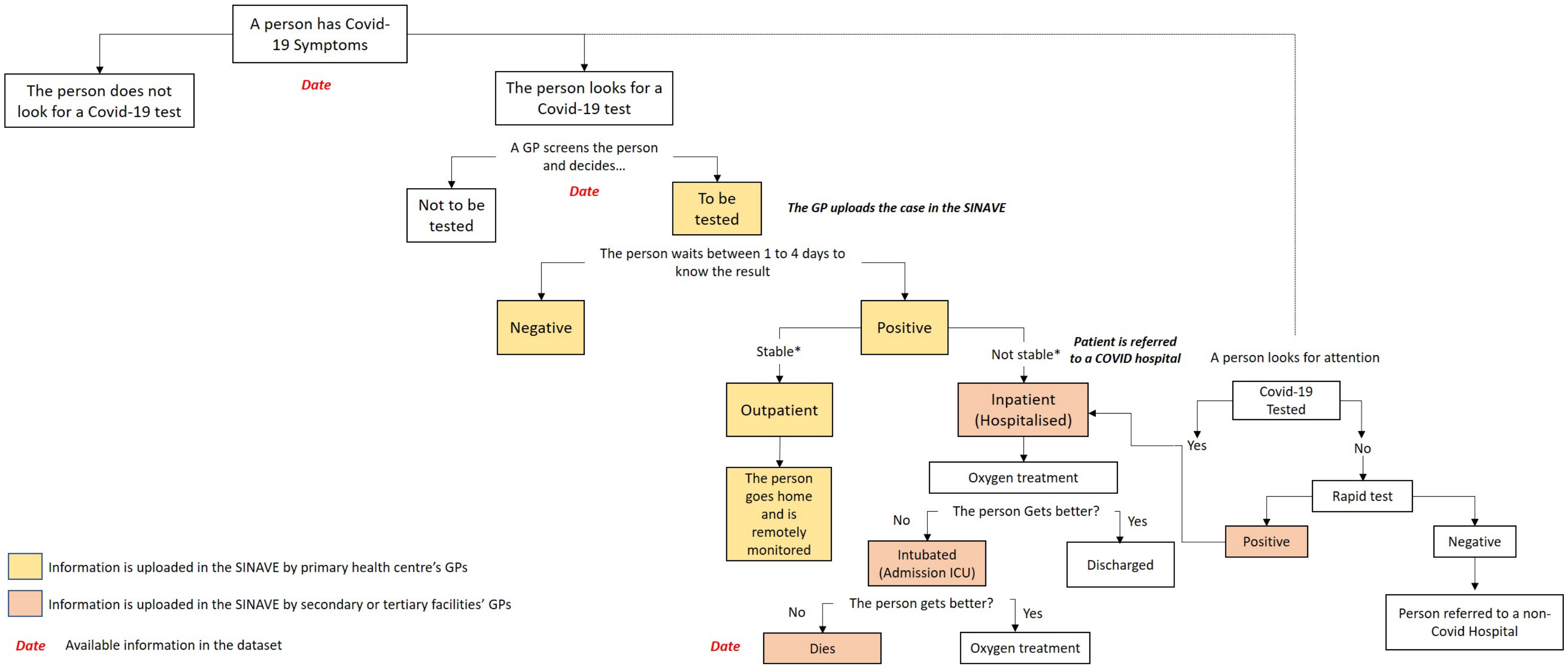
Note: *No Ind. Pop.* means municipalities with no indigenous people; *Dispersed Ind. Pop* means municipalities with a dispersed indigenous population, with less than 40% indigenous population and less than 5,000 indigenous people; *Presence Ind. Pop* means those municipalities with less than 40% indigenous population but more than 5,000 indigenous individuals within its total population; *Indigenous* means municipalities with 70% or more of indigenous population (Comisión Nacional para el Desarrollo de los Pueblos Indigenas, 2015)

C.1 COVID-19 procedures and data collection

The testing procedure was as follows: people who have symptoms and sought out a test arrive at the health unit (this assumes that people are physically capable of going to their regular health unit). Once in the health facility, the general practitioner (GP) screens the patient and decides if the patient meets the inclusion criteria to be tested for COVID-19. If patients are tested, GPs capture information about their medical history, the current date and the date when the patient first showed a symptom. This information is recorded in an online platform called SINAVE (National Epidemiological Surveillance System). Cases in the dataset represent both ambulatory (outpatient) and hospitalised (inpatient) individuals. Swabs are obtained from outpatients and samples are sent to the nearest Laboratory of Respiratory Virus (InDRE). This process could take up to four days. If the case is positive, there are two potential paths to follow which depends on the health status of the patient. If the person is clinically assessed and diagnosed as with a mild to moderate infection, the person can remain at home and be remotely monitored. Follow-up of all suspected COVID-19 cases and ambulatory patients is done by the responsible healthcare professional of every Local Health Jurisdiction. This person is also in charge of uploading the data into SINAVE. Due to collection procedures, a patient who is tested more than once in different jurisdictions and at different points in time may lead to duplicate records as there is no unique identification variable available to identify individual patients.

A patient clinically diagnosed with a complicated to severe infection (when the patient has difficulties with breathing or hypoxemia) is admitted to a specialised COVID-19 hospital. In these hospitals, patients immediately receive drug and oxygen treatment. If patients do not respond favourably to the treatment, they can be admitted to the intensive care unit (ICU). It is also possible that patients who never asked for a test when they first felt symptoms could arrive at a hospital seeking medical attention, without any previous test or clinical record. In this scenario, patients are rapidly screened and if the test is positive the patient is admitted to a Covid-hospital; if not, the patient is referred to another hospital to receive care. In the case of patients that for some reason are already intubated, bronchoalveolar lavage sample is obtained and tested for COVID-19. If an inpatient died due to suspected COVID-19, lung biopsies are obtained from an autopsy. Reporting of deaths is obligatory and must be done in less than 48 hours after occurrence. If patients are not able to give details about their medical history, this is retrieved from records. All these data are undertaken by accredited hospital epidemiologists and uploaded in the SINAVE.

Figure C.1.1: Diagram of COVID-19 procedures and data collection in Mexico



C.2 Variable definitions

Table C.2.1: Definition of individual-level variables

Dimension	Variable	Definition
		Individual-level characteristics
Demographics	Sex	Sex of the individual. 1 if female, 0 male
	Age	Individual years of age
Underlying Health Conditions	Pneumonia	1 if the patient has a diagnosis of pneumonia, 0 otherwise
	Hypertension	1 if the patient has a diagnosis of hypertension, 0 otherwise
	Diabetes	1 if the patient has a diagnosis of diabetes, 0 otherwise
	COPD	Chronic obstructive pulmonary disease. 1 if the person has a diagnosis of a COPD, 0 otherwise
	Asthma	1 if the patient has a diagnosis of asthma, 0 otherwise
	Immunosuppression	1 if the patient has immunosuppression, 0 otherwise
	Renal disease	1 if the patient has a diagnosis of a renal disease, 0 otherwise
	Cardiovascular disease	1 if the patient has a diagnosis of a cardiovascular disease, 0 otherwise
	Other	Other comorbidities
Risky health behaviours	Obesity	To be obese. 1 if the patient has obesity, 0 otherwise. There is no clinical definition available in the dataset
	Smoking	To smoke. 1 if the patient smokes regularly, 0 otherwise
Institution where individuals received medical attention	Testing waiting-time	Number of days the person waited to get tested since the first symptom
	IMSS	Mexican Social Security Institute
	ISSSTE	Civil Service Social Security and Services Institute
	SSA	Health Ministry. SSA hospitals provide health services to people enrolled in the INSABI programme, former known as "Seguro Popular"
	Federal States	Hospitals owned and managed by the Federal States
	PEMEX	Hospitals owned and managed by the state-owned petroleum company "Mexican Petroleum"
	SEDENA	Hospitals owned and managed by the Secretariat of National Defence
	SEMAR	Hospitals owned and managed by the Secretariat of the Navy

C.3 Linear probability models (LPM) and logit regression models

Tables C.3.1 and C.3.2 depict the results from the linear and nonlinear regression models, respectively. There are two columns associated to each outcome, the first shows the coefficients for non-indigenous($G_i = 0$) and second indigenous($G_i = 1$). Table C.3.2 used the a logit function and coefficients are expressed in log-odds.

Table C.3.1: Linear regression results for indigenous and non-indigenous people. All outcomes

	Hosp_Gi=0		**Hosp_Gi=1**		**Dead_Gi=0**		**Dead_Gi=1**	
Demographics								
Age	0.00***	(0.00)	0.00***	(0.00)	0.00***	(0.00)	0.00***	(0.00)
Women	-0.02***	(0.00)	-0.01**	(0.00)	-0.02***	(0.00)	-0.02***	(0.00)
Comorbidities								
COPD	0.08***	(0.00)	0.05***	(0.01)	0.04***	(0.00)	-0.01	(0.01)
Asthma	-0.01***	(0.00)	-0.01	(0.01)	-0.01***	(0.00)	0.00	(0.01)
Immunosuppression	0.10***	(0.00)	0.14***	(0.02)	0.01***	(0.00)	-0.01	(0.01)
Renal D.	0.16***	(0.00)	0.10***	(0.01)	0.08***	(0.00)	0.05***	(0.01)
Pneumonia	0.69***	(0.00)	0.69***	(0.00)	0.32***	(0.00)	0.35***	(0.00)
Other C.	0.08***	(0.00)	0.09***	(0.01)	0.02***	(0.00)	0.03**	(0.01)
NCD								
Diabetes	0.06***	(0.00)	0.07***	(0.01)	0.03***	(0.00)	0.01***	(0.00)
Hypertension	0.03***	(0.00)	0.03***	(0.01)	0.02***	(0.00)	0.02***	(0.00)
Cardio D.	0.07***	(0.00)	0.07***	(0.01)	0.02***	(0.00)	0.00	(0.01)
Risky Behaviours								
Smoking	-0.01***	(0.00)	-0.02*	(0.01)	-0.00***	(0.00)	-0.01	(0.01)
Obesity	-0.00***	(0.00)	-0.00	(0.01)	0.00***	(0.00)	0.02***	(0.00)
Medical Att.								
Wait Test	0.00***	(0.00)	0.00***	(0.00)	0.00***	(0.00)	0.00***	(0.00)
Private	-0.15***	(0.00)	-0.19**	(0.06)	-0.02***	(0.00)	-0.03	(0.05)
IMSS	-0.11***	(0.00)	-0.11+	(0.06)	0.05***	(0.00)	0.05	(0.05)
ISSSTE	-0.11***	(0.00)	-0.12*	(0.06)	0.01**	(0.00)	0.02	(0.05)
SSA	-0.20***	(0.00)	-0.21***	(0.06)	0.00	(0.00)	0.00	(0.05)
State	-0.19***	(0.00)	-0.21***	(0.06)	-0.01***	(0.00)	-0.04	(0.05)
PEMEX	-0.16***	(0.00)	-0.24***	(0.06)	-0.03***	(0.00)	-0.05	(0.05)
SEDENA	0.17***	(0.00)	0.28***	(0.06)	0.03***	(0.00)	0.13*	(0.05)
SEMAR	-0.12***	(0.00)	-0.14+	(0.08)	-0.01*	(0.00)	-0.05	(0.06)
Health Infrastructure								
No Affiliated	0.00***	(0.00)	0.00**	(0.00)	0.00***	(0.00)	0.00	(0.00)
Health Fac.	-0.00***	(0.00)	-0.00***	(0.00)	-0.00	(0.00)	0.00	(0.00)
Urban loc.	-0.00***	(0.00)	-0.00	(0.00)	-0.00***	(0.00)	-0.00	(0.00)
Household Ch.								
Overcrow	0.00	(0.00)	0.00	(0.01)	-0.00+	(0.00)	0.01*	(0.01)
Perc. Floors	-0.00***	(0.00)	-0.00	(0.00)	-0.00+	(0.00)	-0.00	(0.00)
Perc. Sleeping	0.00**	(0.00)	0.00	(0.00)	0.00**	(0.00)	0.00	(0.00)
Perc. Room	-0.01***	(0.00)	-0.01*	(0.00)	-0.00**	(0.00)	0.00	(0.00)
No Water	-0.00**	(0.00)	-0.00+	(0.00)	0.00**	(0.00)	-0.00	(0.00)
No Elec.	-0.04***	(0.00)	-0.03***	(0.01)	-0.01***	(0.00)	-0.01	(0.00)
Latrine	-0.00*	(0.00)	-0.01***	(0.00)	-0.00***	(0.00)	-0.00	(0.00)
No Drainage	0.00***	(0.00)	0.01***	(0.00)	0.00***	(0.00)	0.00	(0.00)
No E,W,D	0.01*	(0.00)	0.05***	(0.01)	-0.01***	(0.00)	-0.01	(0.01)
Perc. No car	0.00***	(0.00)	-0.00	(0.00)	-0.00***	(0.00)	0.00+	(0.00)
Perc. No appl.	0.00	(0.00)	0.03***	(0.01)	-0.01***	(0.00)	-0.01	(0.01)
Perc. No TV	0.01***	(0.00)	0.01	(0.00)	-0.01***	(0.00)	-0.01*	(0.00)
Perc. No phone	-0.01***	(0.00)	-0.01**	(0.00)	-0.00***	(0.00)	-0.01***	(0.00)
Perc. No Comp.	0.00***	(0.00)	0.01***	(0.00)	0.00***	(0.00)	0.00***	(0.00)
Perc. NO ICT	0.02***	(0.00)	-0.01	(0.01)	0.02***	(0.00)	0.02*	(0.01)
Economic Ch.								
Perc. No School Ch	0.02***	(0.00)	0.01	(0.01)	-0.00**	(0.00)	-0.01	(0.01)
Perc. Illiterate Ch.	-0.01***	(0.00)	0.00	(0.01)	0.01***	(0.00)	0.00	(0.01)
Perc. Illiterate Adu.	0.00+	(0.00)	-0.00	(0.00)	-0.00	(0.00)	0.00	(0.00)
Perc. No School a	-0.00***	(0.00)	0.00	(0.00)	-0.00	(0.00)	-0.00	(0.00)
Perc. Unemployed	-0.02***	(0.00)	0.00	(0.00)	-0.00***	(0.00)	-0.01+	(0.00)
_cons	0.15***	(0.00)	0.14*	(0.07)	-0.07***	(0.00)	-0.17**	(0.06)
N	4,765,878		30,930		4,765,878		30,930	
r2	.499		.526		.291		.311	

Notes: standard errors in parenthesis + p<0.1, * p<0.05, ** p<0.01, *** p<0.001

Table C.3.2: Nonlinear regression results for indigenous and non-indigenous people. All outcomes

	Hosp_Gi=0		**Hosp_Gi=1**		**Dead_Gi=0**		**Dead_Gi=1**	
Demographics								
Age	0.03***	(0.00)	0.01***	(0.00)	0.06***	(0.00)	0.05***	(0.00)
Women	-0.37***	(0.00)	-0.12**	(0.04)	-0.54***	(0.01)	-0.39***	(0.05)
Comorbidities								
COPD	0.52***	(0.01)	0.38***	(0.11)	-0.07***	(0.01)	-0.25*	(0.10)
Asthma	-0.22***	(0.01)	-0.10	(0.12)	-0.21***	(0.02)	0.06	(0.14)
Immunosuppression	1.04***	(0.01)	1.21***	(0.15)	0.41***	(0.02)	0.19	(0.18)
Renal D.	1.20***	(0.01)	0.85***	(0.12)	0.64***	(0.01)	0.40***	(0.12)
Pneumonia	4.11***	(0.01)	3.92***	(0.05)	2.81***	(0.01)	2.79***	(0.05)
Other C.	0.75***	(0.01)	0.67***	(0.11)	0.25***	(0.01)	0.39**	(0.13)
NCD								
Diabetes	0.63***	(0.01)	0.61***	(0.05)	0.40***	(0.01)	0.27***	(0.06)
Hypertension	0.30***	(0.01)	0.28***	(0.05)	0.16***	(0.01)	0.17**	(0.06)
Cardio D.	0.55***	(0.01)	0.63***	(0.12)	-0.08***	(0.01)	-0.08	(0.13)
Risky Behaviours								
Smoking	-0.20***	(0.01)	-0.24**	(0.09)	-0.20***	(0.01)	-0.22*	(0.10)
Obesity	0.04***	(0.01)	-0.03	(0.06)	0.32***	(0.01)	0.40***	(0.06)
Medical Att.								
Wait Test	0.03***	(0.00)	0.03***	(0.01)	0.04***	(0.00)	0.04***	(0.01)
Private	-1.54***	(0.03)	-1.62**	(0.56)	-0.59***	(0.06)	-1.40+	(0.77)
IMSS	-0.98***	(0.03)	-0.86	(0.52)	1.15***	(0.06)	-0.05	(0.71)
ISSSTE	-1.07***	(0.03)	-0.94+	(0.53)	0.21***	(0.06)	-0.47	(0.71)
SSA	-2.42***	(0.03)	-1.87***	(0.52)	-0.18**	(0.06)	-0.65	(0.71)
State	-2.14***	(0.03)	-1.82**	(0.58)	0.09	(0.06)	-1.06	(0.77)
PEMEX	-1.70***	(0.04)	-2.14***	(0.62)	-0.34***	(0.06)	-1.40+	(0.80)
SEDENA	0.95***	(0.03)	1.53**	(0.53)	0.80***	(0.06)	0.88	(0.72)
SEMAR	-1.05***	(0.04)	-1.13	(0.71)	-0.01	(0.07)	-1.91	(1.32)
Health Infrastructure								
No Affiliated	0.00***	(0.00)	0.01+	(0.00)	0.01***	(0.00)	0.00	(0.00)
Health Fac.	-0.00***	(0.00)	-0.00**	(0.00)	-0.00***	(0.00)	0.00	(0.00)
Urban loc.	-0.00***	(0.00)	-0.00***	(0.00)	-0.00***	(0.00)	-0.00	(0.00)
Household Ch.								
Overcrow	-0.03*	(0.01)	0.08	(0.08)	-0.10***	(0.02)	0.09	(0.10)
Perc. Floors	-0.06***	(0.01)	-0.02	(0.02)	-0.02*	(0.01)	-0.02	(0.02)
Perc. Sleeping	-0.01***	(0.00)	0.02	(0.02)	-0.03***	(0.00)	0.01	(0.02)
Perc. Room	-0.13***	(0.00)	-0.05*	(0.02)	0.00	(0.01)	0.02	(0.03)
No Water	-0.02***	(0.00)	-0.03*	(0.01)	-0.01*	(0.00)	-0.02	(0.02)
No Elec.	-0.49***	(0.03)	-0.26***	(0.06)	-0.29***	(0.04)	-0.02	(0.08)
Latrine	-0.01***	(0.00)	-0.07***	(0.01)	-0.02***	(0.00)	-0.01	(0.01)
No Drainage	0.06***	(0.00)	0.06***	(0.01)	0.03***	(0.01)	0.00	(0.01)
No E,W,D	0.16**	(0.06)	0.42***	(0.11)	0.04	(0.08)	-0.02	(0.14)
Perc. No car	0.01***	(0.00)	-0.00	(0.01)	-0.00+	(0.00)	0.01	(0.01)
Perc. No appl.	-0.03	(0.03)	0.26**	(0.08)	0.14***	(0.04)	-0.06	(0.10)
Perc. No TV	0.24***	(0.01)	0.07	(0.05)	-0.01	(0.02)	-0.10+	(0.06)
Perc. No phone	-0.05***	(0.01)	-0.07*	(0.03)	-0.03***	(0.01)	-0.11**	(0.04)
Perc. No Comp.	0.02***	(0.00)	0.05***	(0.01)	0.03***	(0.00)	0.04**	(0.01)
Perc. NO ICT	0.10**	(0.04)	-0.09	(0.12)	0.12*	(0.05)	0.27+	(0.16)
Economic Ch.								
Perc. No School Ch	0.26***	(0.02)	0.09	(0.09)	0.01	(0.03)	-0.00	(0.12)
Perc. Illiterate Ch.	-0.00	(0.03)	0.01	(0.08)	0.22***	(0.04)	-0.02	(0.12)
Perc. Illiterate Adu.	0.00	(0.01)	-0.02	(0.02)	-0.06***	(0.01)	0.02	(0.03)
Perc. No School a	-0.02***	(0.01)	0.02	(0.02)	0.02*	(0.01)	-0.01	(0.03)
Perc. Unemployed	-0.31***	(0.01)	-0.02	(0.05)	-0.17***	(0.01)	-0.08	(0.07)
_cons	-2.41***	(0.06)	-2.41***	(0.64)	-7.03***	(0.09)	-6.36***	(0.86)
N	4,765,878		30,930		4,765,878		30,930	
r2_p	.477		.46		.453		.403	

Notes: standard errors in parenthesis + p<0.1, * p<0.05, ** p<0.01, *** p<0.001

C.4 Oaxaca-Blinder decomposition approach using Linear and Nonlinear models

Tables C.4.1 to C.4.3 show the results of the aggregate and detailed Oaxaca decompositions using nonlinear and linear models.

Table C.4.1: Aggregate Oaxaca Decomposition. Linear models

	Hospitalisations	Deaths	%	pctl_y4
Non Indigenous	0.127***		0.050***	
	(0.00)		(0.00)	
Indigenous	0.245***		0.097***	
	(0.00)		(0.00)	
Mean Difference	-0.118***		-0.047***	
	(0.00)		(0.00)	
Explained	-0.093***	79.507***	-0.041***	88.159***
	(0.00)	(1.23)	(0.00)	(2.14)
Unexplained	-0.024***	20.493***	-0.006***	11.841***
	(0.00)	(1.23)	(0.00)	(2.14)
Observations	4,796,808	4,796,808	4,796,808	4,796,808

Notes: Bootstrapped standard errors in parenthesis (500 replications)
Models fitted using an ANOVA-type normalisation and weights from
a first-order Taylor linearisation. % share of each component to the overall gap.
+ p<0.1, * p<0.05, ** p<0.01, *** p<0.001

Table C.4.2: Detailed Oaxaca Decomposition for Hospitalisations. Linear and Nonlinear models

	Nonlinear		Linear	
E. Demographics	-0.006***	(0.00)	-0.004***	(0.00)
%	5.061***	(0.53)	3.567***	(0.36)
E. UndLyingCond	-0.054***	(0.00)	-0.068***	(0.00)
%	46.282***	(1.15)	57.442***	(1.15)
E. Risky Behav.	-0.001*	(0.00)	-0.001*	(0.00)
%	0.931*	(0.36)	0.494*	(0.21)
E. Med. Attent.	0.013***	(0.00)	0.008***	(0.00)
%	-10.683***	(0.95)	-6.469***	(0.75)
E. Health Infra.	-0.008*	(0.00)	-0.004*	(0.00)
%	6.695*	(2.69)	3.773*	(1.49)
E. Household Ch.	-0.039***	(0.01)	-0.026***	(0.00)
%	32.954***	(5.88)	21.791***	(4.11)
E. Mun. Eco. Ch.	-0.001	(0.01)	0.001	(0.00)
%	0.470	(5.25)	-1.093	(3.92)
Ue. Demographics	0.022***	(0.00)	0.016**	(0.01)
%	-19.132***	(2.13)	-13.434**	(4.43)
Ue. UndLyingCond	0.001*	(0.00)	-0.001	(0.00)
%	-0.835*	(0.38)	0.699	(0.92)
Ue. Risky Behav.	0.000	(0.00)	0.001	(0.00)
%	-0.394	(0.34)	-1.097	(0.74)
Ue. Med. Attent.	-0.015	(0.04)	0.006	(0.07)
%	12.453	(35.37)	-5.434	(62.68)
Ue. Health Infra.	0.003	(0.00)	-0.003	(0.01)
%	-2.272	(3.04)	2.506	(6.89)
Ue. Household Ch.	-0.021	(0.02)	-0.029	(0.04)
%	17.974	(13.44)	24.461	(32.53)
Ue. Mun. Eco. Ch.	-0.013***	(0.00)	-0.022**	(0.01)
%	10.655***	(2.39)	18.410**	(6.12)
Intercept	0.000	(0.05)	0.007	(0.08)
%	-0.157	(38.42)	-5.617	(71.68)
Observations	4,796,808		4,796,808	

Notes: Bootstrapped standard errors in parenthesis (500 replications)
Nonlinear models fitted using an ANOVA-type normalisation
and weights from a first-order Taylor linearisation
% share of each component to the overall gap
+ p<0.1, * p<0.05, ** p<0.01, *** p<0.001

Table C.4.3: Detailed Oaxaca Decomposition for Deaths. Linear and nonlinear models

	Nonlinear		Linear	
E. Demographics	-0.014***	(0.00)	-0.008***	(0.00)
%	30.295***	(1.86)	17.464***	(0.86)
E. UndLyingCond	-0.024***	(0.00)	-0.032***	(0.00)
%	50.678***	(2.76)	68.224***	(1.84)
E. Risky Behav.	-0.001*	(0.00)	-0.000	(0.00)
%	1.909*	(0.78)	0.717	(0.53)
E. Med. Attent.	0.004**	(0.00)	0.003***	(0.00)
%	-8.552**	(2.73)	-7.099***	(1.40)
E. Health Infra.	0.001	(0.00)	0.001	(0.00)
%	-2.825	(5.11)	-3.027	(2.92)
E. Household Ch.	-0.002	(0.01)	-0.003	(0.00)
%	5.097	(12.25)	6.587	(7.53)
E. Mun. Eco. Ch.	-0.005	(0.01)	-0.002	(0.00)
%	11.032	(10.76)	5.294	(7.51)
Ue. Demographics	0.005***	(0.00)	-0.021***	(0.00)
%	-10.242***	(2.27)	45.707***	(8.77)
Ue. UndLyingCond	0.000	(0.00)	-0.000	(0.00)
%	-0.276	(0.26)	0.065	(2.39)
Ue. Risky Behav.	-0.000	(0.00)	-0.002+	(0.00)
%	0.171	(0.28)	3.716+	(1.98)
Ue. Med. Attent.	0.007	(0.05)	-0.003	(0.06)
%	-14.691	(99.80)	5.806	(126.97)
Ue. Health Infra.	-0.000	(0.00)	-0.005	(0.01)
%	0.095	(2.35)	9.824	(14.45)
Ue. Household Ch.	-0.010+	(0.01)	-0.087**	(0.03)
%	21.291+	(11.06)	185.661**	(63.66)
Ue. Mun. Eco. Ch.	-0.001	(0.00)	0.003	(0.00)
%	2.631	(2.04)	-6.736	(10.00)
Intercept	-0.006	(0.05)	0.109+	(0.07)
%	13.387	(100.04)	-232.202+	(138.69)
Observations	4,796,808		4,796,808	

Notes: Bootstrapped standard errors in parenthesis (500 replications)
Nonlinear models fitted using an ANOVA-type normalisation
and weights from a first-order Taylor linearisation
% share of each component to the overall gap
+ p<0.1, * p<0.05, ** p<0.01, *** p<0.001