ASSESSMENT AND PREDICTION OF BODY MASS INDEX (BMI) DISTRIBUTIONS AMONG ADULT POPULATIONS IN MEXICO, COLOMBIA, AND PERU, 1988-2014

by

Goro Yamada

A dissertation submitted to Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy

Baltimore, Maryland

October, 2016

Abstract

Unequal distribution of adult obesity across subpopulations in low- and middle-income countries has been reported, but findings have been mainly from data of women of reproductive age. While mean body mass index (BMI), overweight and obesity prevalence are commonly used obesity indicators, incorporation of ever-changing skewed BMI distributions has been a challenge. In this context, our study aimed to assess differences in magnitude and rates of change in BMI distributions by sex, age, geographic and socioeconomic factors in Mexico, Colombia, and Peru by modeling entire BMI distributions. Furthermore, this modeling technique was applied for the prediction of future obesity indicators.

Data from nationally representative health surveys conducted between 1988 and 2014 in these 3 countries were used. The analyses were conducted using the generalized additive model for location, scale, and shape (GAMLSS) in order to model BMI distributions. BMI was assumed to follow a Box-Cox Power Exponential (BCPE) distribution, and each of its 4 parameters was modeled as a function of demographic, geographic, and socioeconomic factors. Prediction models were evaluated using data before the last survey, with the predicted values compared to actual values at the time of the last survey.

Whereas women had more right-shifted and wider BMI distributions than men across the countries in 2010, men generally experienced more rapid increases in BMI between 2005 and 2010. More education was negatively associated with BMI in women after covariate adjustment whereas it was somewhat positively associated in men. Higher household wealth was positively associated with BMI in men. Lower household wealth was associated with higher rates of change in BMI distributions in women. The BCPE-GAMLSS model yielded the best prediction performance among the assessed models in predicting obesity prevalence.

Observed differences in BMI distributions across subpopulations suggest the necessity of tailoring relevant policies and programs to reach target populations. Increases in BMI imply increases in obesity-associated diseases, such as cardiovascular diseases and diabetes, for which preventive and

preparative actions would be urgent. The BCPE-GAMLSS method worked well for estimation and prediction of BMI by modeling its distributions precisely.

Committee of Readers

Advisor

Dr. Lawrence H Moulton, PhD, MS

Professor, Department of International Health

Readers

Dr. Carlos Castillo-Salgado, MD, JD, MPH, DrPH

Professor, Department of Epidemiology

Dr. Jessica C Jones-Smith, PhD, MPH

Associate Professor – Adjunct, Department of International Health

Associate Professor, Department of Health Services & Nutrition Sciences Program

University of Washington School of Public Health

Dr. Keshia M Pollack, PhD, MPH

Associate Professor, Department of Health Policy and Management

Acknowledgement

I would like to express my sincere gratitude to Dr. Moulton, who has been my academic and thesis advisor for the second half of my PhD program. His approach has been kind and gentle, and he keeps track of previous and incoming matters. His statistical advice was enormous and helpful in overcoming statistical problems I encountered during the implementation of this study. Without his generous direction and support, I could not have accomplished this work presented here.

My sincere thanks also go to Dr. Castillo-Salgado and Dr. Jones-Smith, who have been my thesis committee members since this study was initiated.

Dr. Castillo-Salgado provided me various comments and advice from his wide perspective in public health practice and in Latin America. His fundamental questions – what the study questions and results could mean in public health policies and practices – have been of a great guide in implementing this study without losing the main track. He gave me a series of lessons how epidemiology could contribute to professional public health.

Dr. Jones-Smith has been always approachable and considerate. Her expertise in obesity – its determinants, mechanisms, consequences, counteracting strategies and challenges – has been a great strength in conducting this research. She also kindly provided me practical lessons on how manuscripts could be effectively written in an audience-focused manner. I am very fortunate that she was able to be on my thesis committee.

I would also like to thank Dr. Pollack, who kindly accepted to serve as a reader of my thesis. She has been energetic and supportive, and provided me another angle to see research – what are take home messages for public health practitioners and decision makers – from her perspective in health policy and management.

I would like to thank Dr. Stan Becker, Department of Population, Family and Reproductive Health for his suggestions in constructing the BMI prediction model, one of the study components. I would also like to thank Dr. Teresa Shamah Levy and her colleagues at National Institute of Public Health, Mexico. They kindly provided me additional information about survey methods used in conducting the National Health and Nutritional Survey in Mexico. I am also grateful to Mr. Mark Miller at the Joint High Performance Computing Exchange (JHPCE) for his kind and continuous assistance in use of the computer cluster at school.

Lastly but not the least, I would like to thank my family. My two sons, Kenji Ignacio and Hibiki Angel, who may still not know what I was doing, gave me energy and encouragement to continue and complete my study. My final thanks go to my wife, Sofia, who has been always patient, thoughtful and supportive with her incredibly continuous positive attitudes. We faced many things to overcome, and we did. I am finally at the completion point after 5 years of study.

Table of contents

Abbreviations

BCPE	Box-Cox Power Exponential
BMI	Body mass index
CENAN	National Center of Food and Nutrition, Peru [Centro Nacional de Alimentación y Nutrición]
CI	Confidence interval
DALYs	Disability-adjusted life years
DEVAN	Directorate General of Food and Nutritional Surveillance, Peru [Dirección Ejecutiva de Vigilancia Alimentaria y Nutricional]
df	Degrees of freedom
DHS	Demographic and Health Survey
ENAHO-MONIN	National Household Survey, Module for Monitoring of Nutritional Indicators, Peru [Encuesta Nacional de Hogares, Modulo de Monitoreo de Indicadores Nutricionales]
ENDS/ENSIN	Demographic and Health Survey / National Survey on Nutritional Situation, Colombia [Encuesta Nacional de Demografía y Salud / Encuesta Nacional de la Situación Nutricional en Colombia]
ENSANUT	National Health and Nutritional Survey, Mexico [Encuesta Nacional de Salud y Nutrición]
FDI	Foreign direct investment
GAMLSS	Generalized additive model for location, scale and shape
GDP	Gross Domestic Product
GNP	Gross National Product
HDI	Human Development Index
HSE	Health Surveys for England, UK

ICBF	Colombian Institute of Family Welfare [Instituto Colombiano de Bienestar Familiar]
INEI	National Institute of Statistics and Information, Peru [Instituto Nacional de Estadística e Informática]
INSP	National Institute of Public Health, Mexico [Instituto Nacional de Salud Pública]
NCD	Non-communicable disease
NHANES	National Health Examination Survey, US
PAR	Population attributable risk
pp	Percentage point
PSU	Primary sampling unit
SD	Standard deviation
SES	Socioeconomic status
SSU	Secondary sampling unit
TSU	Tertiary sampling unit
UN	United Nations
WHO	World Health Organization
WHS	World Health Survey
YLLs	Years of Life Lost

List of tables

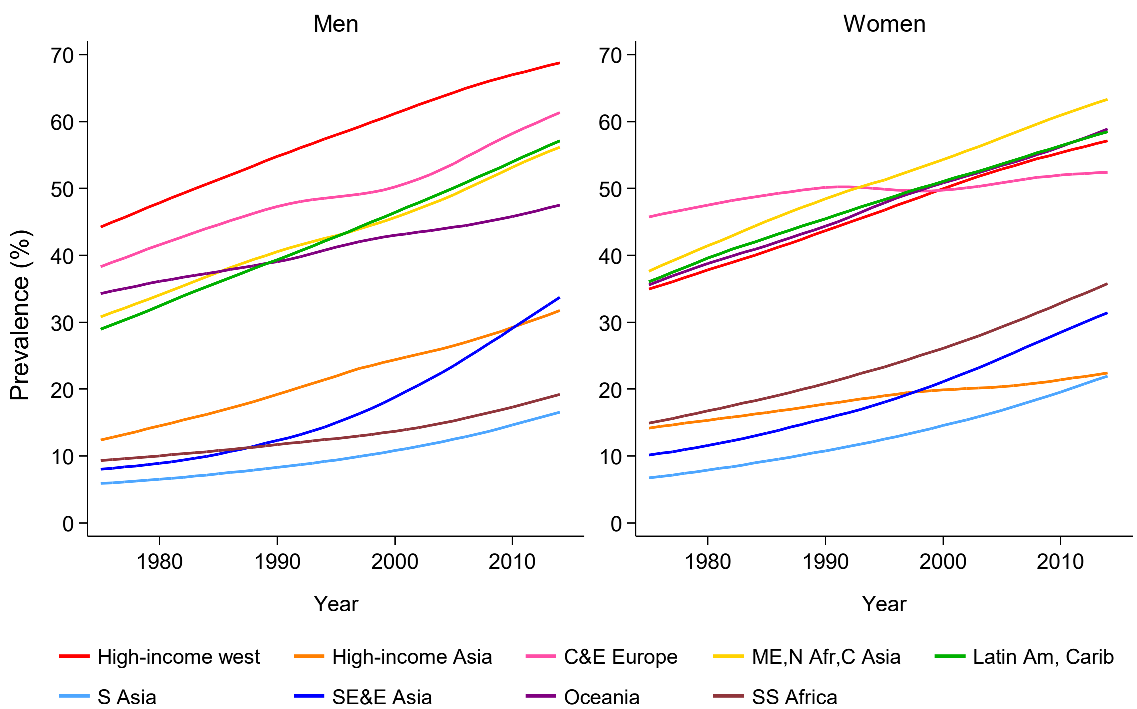
List of figures

Chapter 1 Introduction

1.1 Trends of adult overweight and obesity in low- and middle-income countries in Latin America

Obesity is no longer a problem of only high-income countries. Countries facing obesity problems were mainly in Europe and the United States in the early 20th century. However, obesity started expanding to low- and middle-income countries in recent decades (Caballero 2007). This global epidemic of obesity was formally recognized by World Health Organization (WHO) in the first obesity consultation held in 1997 (WHO 2000, James 2008). Since then, obesity has been recognized as a major public health problem in low- and middle-income countries. Nowadays overweight and obese

Figure 1-1 Trends of age-standardized prevalence of overweight and obesity (BMI $\geq$ 25 kg/m^2) among adult population aged 18 or above, 1975-2014



North Africa, and Central Asia; *Latin Am, Carib*, Latin America and the Caribbean; *S Asia*, South Asia; *SE&E Asia*, Southeast and East Asia; *SS Africa*, Sub-Saharan Africa.

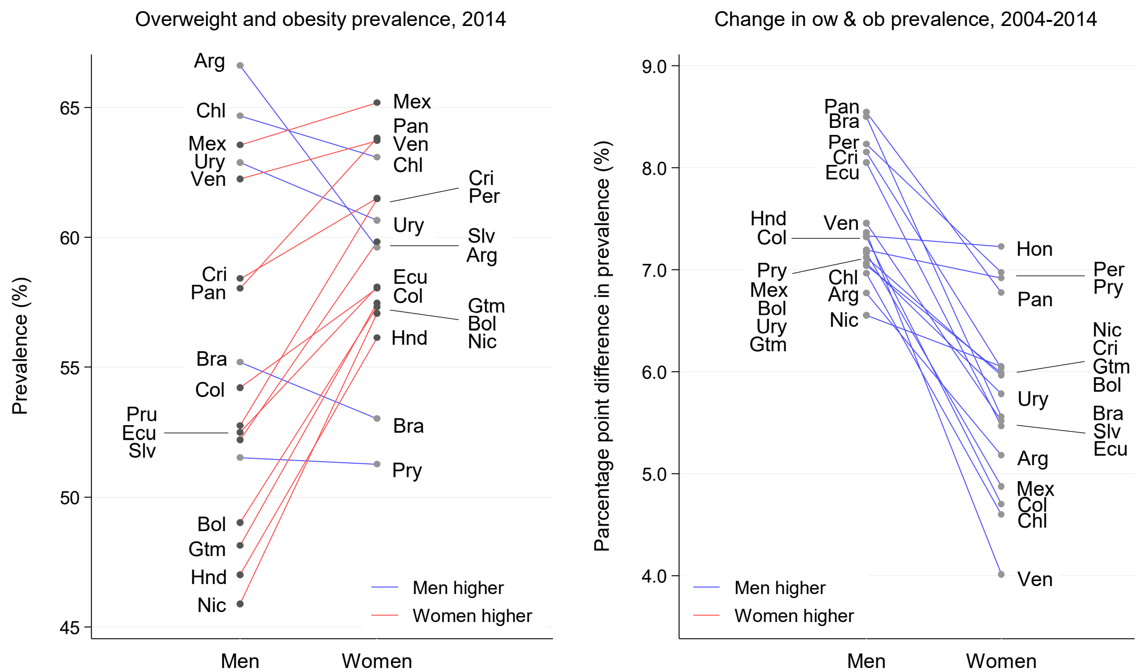
Source: Author's calculation from the data provided by Non-Communicable Disease Risk Factor Collaboration (NCD-RisC). Data obtained from http://www.ncdrisc.org/d-adiposity.html and accessed on April 4, 2016.

populations reside more in low- and middle-income countries. As of 2013, among 1,730 million overweight and obese adults aged 20 years or over worldwide, 1,270 million (73%) were found in low- and middle-income countries (Institute for Health Metrics and Evaluation n.d.).

Between 1975 and 2014, among adult men aged 18 or over, age-standardized overweight and obesity prevalence was the highest in high-income western countries throughout the period (Figure 1-1, NCD Risk Factor Collaboration n.d.). As of 2014, the highest overweight and obesity prevalence was 69% in high-income western countries, followed by Central and Eastern Europe (61%), Latin America and the Caribbean (57%), and the Middle East, North Africa, and Central Asia (56%). However, overweight and obesity prevalence continued to increase in some low- and middle-income countries whereas high-income western countries started to decline in the last decade (5 percentage point (pp) increase in the last decade between 2004 and 2014). These countries included: Southeast and East Asia (11 pp / decade); Central and Eastern Europe; Latin America and the Caribbean; and the Middle East, North Africa, and Central Asia (8 pp / decade in all the 3 regions). On the other hand, among adult women, the Middle East, North Africa, and Central Asia were at the highest in the prevalence throughout the estimated period between 1975 and 2014. As of 2014, overweight and obesity prevalence was the highest in the Middle East, North Africa, and Central Asia (63%), followed by Oceania (59%), Latin America and the Caribbean (59%), and high-income western countries (57%). Rates of change in the prevalence were similar among 7 regions (5-7 pp increase between 2004 and 2014) except Central and Eastern Europe, and high-income Asia where the rates were less.

In Latin America (not including the Caribbean), as of 2014, age-standardized overweight and obesity prevalence ranged widely from 46% to 67% in men, and from 51% to 65% in women (Figure 1-2, NCD Risk Factor Collaboration n.d.). The prevalence tended to be higher in women than men on average; however, the sex differences in prevalence between men and women varied by country. Overweight and obesity prevalence was higher in women than men in 12 (out of 17) countries, among

Figure 1-2 Age-standardized prevalence of overweight and obesity (BMI ≥ 25 kg/m²) in 2014 and their rates of change between 2004-2014 among adults aged 18 years or above in Latin America



ow & ob: overweight and obesity

Arg, Argentina; Bol, Bolivia; Bra, Brazil; Chl, Chile; Col, Colombia; Cri, Costa Rica; Ecu, Ecuador; Gtm, Guatemala; Hnd, Honduras; Mex, Mexico; Nic, Nicaragua; Svl, El Salvador; Pan, Panama; Per, Peru; Pry, Paraguay; Ury, Uruguay; Ven, Venezuela.

Source: Author's calculation from the data provided by Non-Communicable Disease Risk Factor Collaboration (NCD-RisC). Data obtained at http://www.ncdrisc.org/d-adiposity.html and accessed on April 4, 2016.

which 8 countries had more than 5 pp difference (Figure 1-2, left). On the other hand, the rates of change in overweight and obesity prevalence between 2004 and 2014 were higher in men than women in all the countries (Figure 1-2, right). The rates of increase ranged from 6.6 pp to 8.6 in men, and from 4.0 pp to 7.2 pp in women. These rates were higher in almost all the countries than the rate of high-income western countries (5.5 pp / decade in men, and 4.8 pp / decade in women). Sex differences in change of overweight and obesity prevalence between 2004 and 2014 ranged from 0.1 pp to 3.6 pp.

The above region- and country-specific estimates were estimated from models that were developed to provide global estimates for all the countries in the world. Apart from these estimates, the Latin American Consortium of Studies in Obesity (LASO) estimated age-specific and overall obesity prevalence for Latin America by pooling 11 case-control and cohort studies from 8 countries conducted in 1998-2007 (Miranda et al. 2013). Their estimates for overall obesity prevalence among adult populations aged 20 years or over was 14% in men and 18% in women. The Cardiovascular Risk Factor Multiple Evaluation in Latin America (CARMELA) study reported obesity prevalence estimates among adult populations aged 25-64 years by pooling cross-sectional population studies from 7 Latin American cities, which ranged 10-32% in men and 17-30% in women across age groups (Schargrodsky et al. 2008). However, these reports used studies conducted at different years and/or were based on sub-national populations.

1.2 Consequences of overweight and obesity

Overweight and obesity are a major risk factor of various chronic diseases, such as type-2 diabetes, cardiovascular diseases, respiratory-breathing abnormality, and some cancers (Kopelman 2007). In 2010, overweight and obesity were estimated to cause 3.4 million deaths, 4% of years of life lost (YLLs), and 4% of disability-adjusted life-years (DALYs) worldwide (Ng et al. 2014). Overweight and obesity cause diseases through the increased mass of adipose tissue and the increased secretion of pathogenic products from enlarged fat cells (Bray 2004). The increased mass of adipose tissue leads to: *sleep apnea* possibly by obstructing the pharyngeal area; *osteoarthritis* by damaging the joints and tissues with excess weight; and *psychological dysfunction* by stigma against overweight and obesity. The increased secretion of pathogenic products from fat cells leads to: *type-2 diabetes* by causing insulin resistance; *hypertension* and *cardiovascular diseases* by possibly alternating endocrines and their subsequent interactions with organs and tissues. Other related diseases by this mechanism include *cancer* (colon, rectum, and prostate cancer in men; cancer of reproductive system, gallbladder, and breast in women) and *liver and gallbladder diseases* (Bray 2004, Kopelman 2007, Labarthe 2011).

The burden of these diseases is enormous for both people and health systems. People who suffer from these diseases live with chronic disability as well as with the necessity of continuous visits to medical facilities for control and treatment purposes. Besides, in low- and middle-income countries, detection of these non-communicable diseases (NCDs) tends to be delayed due to unrecognized disease progress and limited access to quality health services. As a result, patients die without medical attention or are often required to have more intensive and expensive treatment at hospitals, which can lead to large out of pocket payment because of lack of health insurance (WHO 2011).

On the other hand, treating these NCD's could easily overwhelm the current health care system. Diabetes, one of the major consequences of obesity, is a disease that affects morbidity rather than mortality. The management of diabetes requires a long-term continuous supply of drugs, laboratory exams, and trained health care staff, which can overwhelm existing health systems in low- and middle-income countries that are better prepared for acute illnesses (Beran 2015). The Latin American Consortium of Studies in Obesity (LASO) estimated that the prevalence of diabetes in Latin America and the Caribbean was about 5% according to their study in urban and non-urban samples from 8 countries conducted between 1999 and 2004 (Miranda et al. 2013). International Diabetes Federation (IDF) estimated that the number of adults with diabetes would increase from 29.6 million (prevalence of 9%) in 2015 to 48.8 million (prevalence of 12%) in 2035, with a 65% increase in 25 years (IDF 2014).

Among Latin American populations, obesity seems to contribute more to the incidence of cardiovascular diseases as compared to other populations. The INTERHEART study, a standardized case-control study of acute myocardial infraction in 52 countries, showed abdominal obesity was the most important risk factor among other risk factors, with an average population attributable risk (PAR) of 48.5% (Lanas et al. 2007). The INTERSTROKE study, a standardized case-control study in 32 countries of stroke, showed a moderate association of abdominal obesity with all stroke incidence, with its PAR of 21.3%, which was about a half of the PAR for hypertension, a leading factor (O'Donnell et al. 2016).

1.3 Driving forces of overweight and obesity

Increase of body adiposity is caused by an unequal energy balance caused by excessive dietary energy intake relative to energy expenditure (Caballero 2007). Major driving forces of unbalanced energy levels include: nutrition transition; increasing sedentary behavior and physical inactivity; urbanization; technological improvement; and aging population, connected with globalization and economic growth (Caballero 2007, Swinburn et al. 2011, Popkin et al. 2012) although researchers may relate them somewhat differently.

Diet patterns have changed passing a sequence of stages in human history (Popkin 1994). Diet has changed from a traditional diet high in grains and legumes to the western diet rich in refined carbohydrates, fats, and animal products (Popkin et al. 2012). Among various dietary changes, those that have contributed most to the increases of overweight and obesity include availability of low-cost vegetable oil, availability of caloric sweeteners (especially added sugar in beverages), and increased intake of animal-source food (e.g., meat, eggs, and daily products) (Popkin et al. 2012). In addition, the supply of low-cost, palatable, and energy-dense foods, which can even stimulate addictive eating behavior, became available and consumed thanks to advanced food science research, improved distribution systems, and persuasive marketing (Caballero 2007, Swinburn et al. 2011, Popkin et al. 2012).

Levels of physical activity have reduced over time (Caballero 2007, Popkin et al. 2012). The labor force structure has changed from agricultural labor intensive work, such as farming, mining, and forestry, to sedentary work, such as service and manufacturing industries (Popkin et al. 2012), and work in informal sectors, such as street venders (Fraser 2005). Leisure activities involve more sedentary behavior these days, such as watching television and internet (Caballero 2007, Popkin et al. 2012). Increased mechanization has also led to the decline in physical activity levels (Caballero 2007, Popkin et al. 2012).

More people move to urban areas for better work opportunities and life conditions, and as a consequence, more people are exposed to energy-dense food sold on the street, in fast food restaurants, and in supermarkets (Caballero 2007, Popkin et al. 2012). Consumption of such foods may be accelerated by less time for cooking (Cuevas et al. 2009). Urbanization brought changes in the labor force structure, leisure activities, transportation (Caballero 2007, Cuevas et al. 2009) that led to reduction in levels of physical activity as mentioned previously.

Technological improvement facilitated the increase of mechanization, motorization, and computerization, which in turn caused an increase of sedentary lifestyles (Caballero 2007, Swinburn et al. 2011). The technological improvement also facilitated massive production of low-cost energy-dense food, which eventually led to overconsumption of such food products and obesity (Philipson and Posner 2003, Finkelstein et al. 2005).

Aging facilitates the increase the number of overweight and obesity population and prevalence as well. Because of improvements in living conditions and health care services and systems, more people survive to older ages, including those overweight and obese (Tucker and Buranapin 2001, Leslie and Hankey 2015). In addition, disability increases with age due to the development of chronic conditions, which prevents people from carrying out major activities (Meydani 2001) and could reduce energy expenditure. Obese seniors are less active, which could accelerate weight gain (Jenkins and Fultz 2008), and further decrease physical activity levels (Chan et al. 2013).

Economic development is a precondition for people to develop obesity. At the same time, it brings other changes that include epidemiological transition, aging, urbanization, technological improvement, and nutritional transition (Swinburn et al. 2011), most of which were mentioned as driving forces above. Economic development has facilitated improvement of population health and prolongation of life expectancy by improvement of living environments including health services. Job opportunities have increased especially in urban areas, which attract people from rural areas. Economic development certainly has facilitated the technological improvement, which led in turn to motorization and mechanization, and has enabled the production of cheap and energy-rich food.

In summary, economic development, technological improvement, massive production of cheap and energy-dense food, urbanization, and aging, all interacting with each other, act as direct and indirect causes of excessive amounts of food consumption and/or reduction in levels of physical activity.

1.4 Studies of factors associated with overweight and obesity in low- and middle-income countries

It is now well-known that obesity and obesity-associated diseases are distributed unevenly both across and within low- and middle-income countries. Key factors associated with overweight and obesity that have been investigated include national development (measured by Gross Domestic Product, GDP, Gross National Product, GNP, or Human Development Index, HDI), socioeconomic factors (measured by household income, wealth index, and/or individual educational attainment), and urbanization (measured by place of residence, urban or rural, or proportion of populations in urban areas). Findings from major multi-country studies are summarized in Appendix A.

As a country develops, obesity begins to increase when the country is at a low stage of development, and this association between development and obesity tends to level off as the country develops further. An ecological study of 175 countries showed that, in the year 2007, GDP was positively associated with mean population BMI (men and women combined) until about US$3,000 per capita, then the association became almost flat and insignificant (Egger et al. 2012). Another ecological study with 69 countries demonstrated similar results that population mean BMI increased most rapidly until GDP of about I$5,000 per capita (international dollars), and peaked at about I$17,000 for men and I$12,500 for women according to the data around year 2000 (Ezzati et al. 2005). Other studies demonstrated similar results in low- and middle-income countries among men and women using the data of the World Health Survey (WHS) conducted in 2002-2003 (Nandi et al. 2014) and among women of reproductive age using the data of the Demographic and Health Survey (DHS) and other national surveys (Martorell et al. 2000, Monteiro et al. 2004a, Goryakin and Suhrcke 2014). Furthermore, it was reported that the positive association between GDP/GNP and BMI was weaker in urban areas than those in rural among women (Van Hook et al. 2013, Neuman et al. 2014), and that this positive association reversed in the highest wealth group (Neuman et al. 2014).

Socioeconomic status is associated with overweight and obesity similar to the association observed for the country's developmental status as described above. Numerous studies support that socioeconomic status is positively associated with overweight and obesity when the country's developmental stage is low, and the association starts to level off and becomes negative as the country develops further. In addition, this weakening association seems to start earlier in women than men. A systematic review about the association of socioeconomic status with overweight and obesity was first published in 1989, in which the authors reported a strong positive association of socioeconomic factors with obesity among both men and women in low- and middle-income countries whereas a strong inverse negative association was observed among women and an inconsistent association among men in high-income countries (Sobal and Stunkard 1989). Since then, a series of studies demonstrated that the patterns observed in high-income countries were emerging in low- and middle-income countries as they developed. It was shown that positive associations of socioeconomic factors with obesity were concentrated in countries with low GNP but not in those with high GNP among women of reproductive age (Martorell et al. 2000). A following systematic review demonstrated that obesity was shifting from high to low socioeconomic group in both sexes as GNP increased (Monteiro et al. 2004b). The same study team also showed that a positive association between education and obesity was found in low-income countries whereas a reverse association was observed in upper-middle income countries among women of reproductive age. Later on, various studies confirmed and updated similar associations using data of DHS, WHS, or other sources (Ezzati et al. 2005, Mendez et al. 2005, Subramanian et al. 2011, Pampel et al. 2012, Fleischer et al. 2012, Van Hook et al. 2013, Goryakin and Suhrcke 2014, Aitsi-Selmi et al. 2014) and by systematic reviews (McLaren 2007, Dinsa et al. 2012). Observations that the change of the association between country's development and obesity occurred earlier in women than men were reported by systematic reviews (McLaren 2007, Dinsa et al. 2012) and by data analyses of WHS (Pampel et al. 2012, Fleischer et al. 2012). There were also research findings that positive associations between education and obesity were weaker in more urbanized countries (Fleisher et al. 2012).

Urbanization is generally associated with high BMI, supported by the possible pathways described earlier. This positive association has been reported by various studies (Mendez et al. 2005, Ezzati et al. 2005, Subramanian et al. 2011, Neuman et al. 2013, Nandi et al. 2014, Goryakin and Suhrcke 2014). However, it was also reported that the association might be weakening as a country develops. Among analyzed 42 low- and middle-income countries, little differences were observed in higher GDP countries whereas overweight and obesity prevalence was higher among women in urban areas than rural in lower GDP countries (Popkin et al. 2012). Some other studies implied that the observed positive associations between urban residence and overweight and obesity were explained, to a great extent, by the fact that individuals with higher socioeconomic status reside in urban areas (Subramanian et al. 2011, Neuman et al. 2013).

The above findings about the associations of socioeconomic status and urban residence with BMI were from cross-sectional studies of multiple countries measured at one time-point, and there were uncertainties whether countries would follow the associations and the trajectories implied by one-time-point data of countries at various developmental stages. However, there have been supporting findings from studies conducted using 2 time-points of data. Although people with high socioeconomic status tend to be with higher BMI than those with low socioeconomic status, the rate of increase in BMI seems high among low socioeconomic groups as compared to high socioeconomic groups especially in middle-income countries. A multi-country study demonstrated that high household wealth was associated with lower gains in overweight prevalence among women in 10 out of 37 low- and middle-income countries (Jones-Smith et al. 2011). Another study among ever-married women reported similar findings (Neuman et al. 2011). A subsequent study by Jones-Smith at al. (2012) reported that: increases in overweight prevalence were greater among women in the lowest education group than those in the highest in 27 out of 39 countries; overweight prevalences were greater among women in the lowest wealth group than those in the highest in 11 countries; and higher GDP per capita was associated with more rapid increase in overweight and obesity prevalence in the lowest wealth group as compared to the highest group. Another 2 time-point study reported that, whereas higher socioeconomic factors and urban residence were associated with increases of overweight, the risks of high education, wealth and urban residence were weakening, according to their pooled analyses of 36 low- and middle-income countries (Mamun and Finlay 2014). It was also reported that women working in agriculture and production experienced the largest increases in overweight and obesity while women in higher occupational classes had higher overweight and obesity prevalence in Latin America and the Caribbean, South and southeast Asia, and sub-Saharan African regions (Lopez-Arana et al. 2014).

While these observed associations in low- and middle-income countries are consistent, logical and convincing, the majority of the results come from studies based on data of women of reproductive age (Appendix A). The Demographic and Health Surveys (DHS), which was developed to monitor and evaluate reproductive and child health, are often the only source of population-representative data that contain anthropometric measurements in low- and middle-income countries (Popkin et al. 2012). For example, all the studies that assessed the change in overweight and obesity distributions were based on DHS data (Appendix A). Although some countries started to expand study populations beyond this category, this is a recent movement, and sufficient data are not yet available for inter-country analyses.

The obesity distribution patterns are known to differ by sex and age (Kanter and Caballero 2012), and therefore, observed patterns among women of reproductive age do not necessarily reflect that of the entire population (Popkins et al. 2012). There exist different biological mechanisms for the adiposity distribution by age and sex, such as its hormonal change in women before and after menopause. Sociocultural gender differences also exist, resulting in different food consumption patterns between men and women (Kanter and Caballero 2012).

In Latin America, only some low- and middle-income countries recently conducted the DHS; other countries implement their own national health surveys. In the latter countries, data are kept at each responsible statistical or health national institute, which makes access to data more difficult. Hence, previous studies might not have captured recent obesity trends in Latin America fully.

1.5 Analytical methods of BMI for estimation and prediction

Obesity is defined as an excess of body adiposity. Since it cannot be easily measured in routine examinations, body mass index, body weight (in kilogram) divided by squared height (in meter^2) has been commonly used because of its simplicity in measurement, and its availability and prevalent use in many studies (Philip and Posner 2003, Caballero 2007, Finucane et al. 2011). Other anthropometric measurements, such as waist circumference and waist-to-hip ratio, were thought to be more specific indicators of visceral fat accumulation, adverse metabolic profile and disease risk (Schneider et al. 2010). Furthermore, there is one study that found that the predictive capacity of cardiovascular events and mortality were similar (Emerging Risk Factor Collaboration et al. 2011), and other studies indicated that the central adiposity and BMI predict mortality risk independently (Pischon et al. 2008, Finucane et al. 2011).

Mean BMI, prevalence of overweight and obesity are commonly used to describe historical and future trends of obesity. Whereas these indicators provide a convenient way to capture overall features of obesity situation, a single parameter may not capture the complex movement of BMI distributions. From the US and UK, shapes of the BMI distributions – both its mean and skewness – have been reported moving upwards (Hill et al. 2003, Arterburn et al. 2004, Flegal et al. 2012; Howel 2011, Sperrin et al. 2014). An analysis with quantile-quantile plots of historical BMI data of women from 37 low- and middle-income countries demonstrated that gains in BMI were occurring disproportionately – the higher ends of BMI distributions were increasing more rapidly (Razak et al. 2013). With a similar method, pronounced changes at higher percentiles in the US were reported as well (Krishana et al. 2015). These observations implied that the use of single parameters, such as mean BMI or overweight and obesity prevalence, may not describe well these disproportional changes of the BMI distributions.

Discussion about the importance of modeling the shape of the BMI distributions started recently in order to handle skewed distributions of BMI and their disproportional changes more appropriately and explicitly (Majer et al. 2013, Sperrin et al. 2014). Although previous studies suggested the use of the log-normal distribution to describe skewed BMI distributions (Penman and Jonson 2006, Barendregt and Veerman 2010), it tended to overestimate the proportion of higher BMI values (Majer et al. 2013). For this, the use of the Box-Cox power exponential (BCPE) distribution was suggested, which uses 4 parameters that can be interpreted as relating to median, coefficient of variation, skewness and kurtosis. The BCPE distribution fitted well to the BMI distributions among Dutch population (Rigby and Stasinopoulos 2004, Majer et al. 2013). In fact, the BCPE distribution has been also used to construct international child growth standard curves for weight, length/height, and BMI (WHO 2006). Another study about historical trends of BMI in UK used a latent class analysis, in which multiple latent subpopulations were modeled. They identified two classes of subpopulation: one class was a normal BMI subpopulation who did not gain BMI greatly; and the other was a high BMI subpopulation who tended to gain BMI rapidly (Sperrin et al. 2014).

Up to now, many of the prediction studies for adult BMI used forward extrapolation of historical trends in mean BMI, overweight and obesity prevalence, or the BMI percentiles (Appendix B). Wang et al. (2007) used linear regression to predict mean BMI with the data of the US National Health and Nutrition Examination Surveys (NHANES). Age, year, and their interaction terms were included as covariates, and models were constructed separately for each sex-race group. In the following study, the authors used linear regression models to predict the prevalence of overweight and obesity including year as covariate, for each sociodemographic group. In another component of their study, they attempted to predict future BMI distributions for the entire population incorporating the change in shape of BMI distributions they observed. To do so, they predicted mean BMI for each percentile using linear regression with year as a covariate (Wang et al. 2008). Zaninotto et al. (2009) predicted obesity prevalence using data of the UK Health Survey for England (HSE). The authors calculated sex-, age-, social class-specific obesity prevalence for each year, and then, they fitted linear, power, and exponential curves on the estimated prevalence to predict future prevalence. Ruhm (2007) predicted future overweight and obesity prevalence in the US using quantile regression. The author modeled each BMI percentile with race, age and year as covariates using the NHANES data. Mills (2009) used compositional data analysis, where forecasted proportions remain positive and they sum up to 1, and modeled prevalence of 3 BMI categories directly. Groups of researchers provided global estimates (i.e., country- and regional-level estimates for all countries and regions) for mean BMI, overweight and obesity prevalence from the pooled data of population-based studies from countries in the world. The authors used Bayesian hierarchical models in which estimates were informed by the country's data, and data in other countries, especially those in the same region in similar time periods (Finucane et al. 2011, Stevens et al. 2012, Ng et al. 2014, NCD Risk Factor Collaboration 2016).

Other studies applied microsimulation methods to predict future overweight and obesity distributions (Appendix B). Studies with the UK Foresight model used a 2-step approach. First, they fitted categorical regression models on the historical cross-sectional HSE data to construct prediction models for future prevalence of BMI categories, including age and time as covariates for each sex, and predicted future age-, sex-, and time-specific prevalence. Second, they applied microsimulation models, in which future vital status, disease-status, and BMI of virtual populations were predicted from random draws with assumed disease incidence and mortality rates, and the probabilities of being in each BMI category calculated from the first model (McPherson et al. 2007, Wang et al. 2011, Webber et al. 2012). Basu (2010) calculated annual transition probabilities between 5 BMI categories from longitudinal survey data and applied a simulation technique for prediction.

Majer et al. (2013) predicted the entire BMI distributions for the Dutch population. The authors first estimated 4 parameters of the Box-Cox Power Exponential (BCPE) distribution for each combination of sex, age and survey year. Then, they constructed prediction models for each of the 4 BCPE distribution parameters as a function of age and time using the Lee-Carter model (Lee and Carter 1992). Finally, from the predicted parameters, the BMI distributions were constructed. This method was advantageous in a sense that the BCPE fits well with non-normal skewed BMI distributions and that obesity indicators (median BMI, overweight and obesity prevalence) could be easily calculated from the estimated BMI density functions.

Prediction methods, in which each percentile of BMI distributions is modeled (Ruhm 2007, Wang et al. 2008) or the entire distributions are modeled (Majer et al. 2013), incorporate disproportional changes or differences in BMI distributions. However, according to our best knowledge, how these models would better perform than the methods in which obesity indicators (e.g., mean/median BMI, overweight and obesity prevalence) are modeled directly has not been reported.

1.6 Justification

Men and women are different biologically and behaviorally, and it has been acknowledged that generalizability of study findings from women of reproductive age to the general population is limited. Our study attempted to describe trends of BMI distributions and assess associated geographic and socioeconomic factors using nationally representative surveys including both men and women. This examination is now possible because some Latin American countries started implementing their own national health and nutritional surveys, beyond the issues of maternal and child health, for their entire populations.

Shapes of BMI distributions do not change in a proportional manner. Modeling the entire BMI distributions could provide a more accurate picture and trends of the obesity situations than assessments using representative obesity indicators such as mean BMI, overweight and obesity prevalence. Our study intended to analyze historical trends of obesity, assess the association of geographic and socioeconomic factors with obesity, and predict future BMI distributions by modeling the entire BMI distributions.

Prediction of future obesity trends is a prerequisite step for policy review and strategic planning in public health. For example, number of overweight and obesity cases provides a basis for selection of interventions and target groups. Number of future cases and prevalence of obesity-associated diseases (e.g., cardiovascular diseases, hypertension, and diabetes) could be estimated, from which future scenarios and costs could be simulated in order to modify or formulate strategic plans for future. In addition, due to the long latency period of these diseases, obesity-associated conditions and diseases cannot be solved in a short period of time. Hence, such estimation and prediction would be indispensable to avoid a situation that it is too late to tackle the problems effectively.

Whereas sophisticated modeling methods (e.g., Bayesian hierarchical models, simulation models) became available, their applicability to practical public health settings may be still unfeasible in low- and middle-income countries due to their complicated assumptions and required number of data sources. Hence, it was necessary to find methods that would be methodologically robust and feasible to implement.

1.7 Selection of countries

In order to conduct this study, first, data of nationally representative surveys were searched referring to the list of data sources provided by Ng et al. (2014), from which the authors calculated global estimates of overweight and obesity prevalence. They identified data from major multi-country survey programs (e.g., DHS, World Health Surveys), 3 large databases (i.e., WHO Global Infobase, International Association for the Study of Obesity Data Portal, and Global Health Data Exchange), and scientific literatures published from 1980 to 2012. Second, the data sources were searched by independently visiting the above-mentioned large databases, World Bank Microdata Library, and the websites of ministries of health and national statistics offices in each Latin American country.

Eligible surveys were those that collected individual weight and height measurements from the adult populations of both sexes in Latin America, and whose microdata were available publicly. Initially, countries that conducted such surveys at least at 2 time points were considered. The Caribbean countries were not included since their population characteristics and living environments were thought to be quite distinct from those on the continent. These first criteria identified Argentina, Chile, Colombia, Brazil, Mexico, and Peru. Then, Argentina was excluded since weight and height were self-reported. Brazil was also excluded since one of the surveys did not have appropriate sampling weights. Finally, Colombia, Mexico, and Peru were selected for the study since they conducted surveys at least 4 time points, which was necessary for one of the three study objectives. These 3 countries were classified as upper middle-income countries as of 2010, whose GNI per capita were: US$8,650 (Mexico); 5,540 (Colombia); and 4,380 (Peru) (Atlas method, current US$) (World Bank 2016a). For the year 2010, countries were classified as upper middle-income countries if their GNI per capita were between US$3,976-12,275 (World Bank 2016b).

1.8 Ethical approval and funding

This study was reviewed by the Johns Hopkins Bloomberg School of Public Health Institutional Review Board on November 6, 2014. The study was secondary analyses of existing, de-identified data and de-linked databases, and was determined to be exempt. Thus, continuing review of this study as human subjects research was unnecessary. No funding was received to conduct this study.

Chapter 2 Magnitude and rates of change in BMI distributions by age and sex

2.1 Objective

The objective of this component of the study is to assess differences in magnitude and rates of change in BMI distributions by age in both sexes.

2.2 Methods

2.2.1 Data sources

Data at two time points from nationally representative health and nutritional household surveys conducted between 2005 and 2013 in Mexico, Colombia, and Peru were used. All the surveys were household surveys implemented with stratified multistage cluster sampling. Details about sampling designs of each survey are summarized in Appendix C.

For Mexico, data from the National Health and Nutritional Survey (Encuesta Nacional de Salud y Nutrición, ENSANUT) conducted in 2006 and 2012 were used. ENSANUT aimed: to quantify health and nutritional conditions of the population, and their determinants; and to assess health services and programs that attempted to improve them (Olaiz-Fernández et al. 2006, Shamah-Levy et al. 2007, Gutiérrez et al. 2013). From selected households, household members were randomly selected from each defined group (children under 5 years of age, children aged 5-9 years, adolescents, adults, and service users). The dataset was obtained from the website of the National Public Health Institute, Mexico, after authorization (INSP 2009a, 2009b, 2012a, 2012b).

For Colombia, data from the Demographic and Health Survey / Nutritional Situation National Household Surveys (Encuesta Nacional de Demografía y Salud / Encuesta Nacional de la Situación Nutricional en Colombia, ENDS/ENSIN) conducted in 2005 and 2010 were used. ENDS aimed to monitor demographic and health indicators especially related with maternal and child health (Profamilia and Macro International 2005a, Profamilia and IFC Macro 2011a). ENSIN aimed to estimate prevalence of major nutritional problems and their determinants (ICBF 2006, 2011). ENDS and ENSIN had each data collection team in 2005. ENSIN was conducted with a subsample of ENDS, but anthropometry data were collected by ENDS from the full sample. In 2010, ENDS and ENSIS unified data collection teams. In both surveys in 2005 and 2010, they collected anthropometric information from all household members aged 0 to 64. Datasets were obtained from the archive of the DHS Program after authorization (Profamilia and Macro International 2005b, Profamilia and IFC Macro 2011b).

For Peru, data from the National Household Survey, Module for Monitoring of Nutritional Indicators (Encuesta Nacional de Hogares, Modulo de Monitoreo de Indicadores Nutricionales, ENAHO-MONIN) conducted in 2007-8 and 2012-3 were used. ENAHO aimed to provide indicators about living condition, poverty, and impact of social programs (INEI 2012a, 2012b, 2013a). MONIN was an attached module to ENAHO, which aimed to monitor anthropometric indicators and hemoglobin status of the population (CENAN and INEI 2009, DEVAN 2015a). MONIN surveys visited subsamples of the households selected for the preceding main ENAHO surveys conducted at specific semesters. They collected information from all household members. The ENAHO data and the MONIN 2007-8 data were obtained from the website of the National Institute of Statistics and Information, Peru (INEI 2012c, 2012d, 2013b, INEI and CENAN 2010b), and the MONIN 2012-3 data were obtained from the websites of the National Institute of Health, Peru (DEVAN 2015b).

In all 3 countries, primary sampling units (PSUs) were sampled from each stratum within each sub-national unit (state in Mexico and department in Colombia and Peru) with probability proportional to size of dwellings or populations. A stratification unit consisted of a combination of the sub-national unit and the urban-rural area or finer area category. The ultimate sampling unit was dwelling, from which all households were sampled. All household members (in Colombia and Peru) or some of the household members (in Mexico) were followed with household and individual questionnaires. The selected household members were measured for anthropometry, and their blood specimens were collected. Weight and height were measured using preset standardized procedures by trained survey teams.

Country-level household response rates ranged from 85 to 92% (Table 2-1). The rates were lower in urban area (81-90%) than rural area (89-95%) in Colombia and Peru where data were available. Overall individual response rate for obtaining anthropometric measurements ranged from 75 to 84%

Table 2-1 Household and anthropometry response rate

	Mexico		Colombia		Peru	
	2006	2012	2005	2010	2007-8	2012-3
Household response rate*						
Urban	--	--	87%	90%	87%	≈ 81%
Rural	--	--	92%	95%	89%	≈ 90%
Total	--	87%	88%	92%	87%	≈ 85%
Anthropometry response rate† (% not responding due to absence‡)						
Men	67% (--)	81% (--)	61% (36%)	72% (25%)	79% (13%)	68% (17%)
Women	80% (--)	84% (--)	86% (12%)	88% (10%)	89% (6%)	81% (10%)
Total	75% (--)	83% (--)	75% (23%)	80% (17%)	84% (9%)	75% (13%)

-- Not available

* Household response rate: as reported in a publication (Mexico. Perez-Escamilla, et al. 2014); calculated as the number of interviewed households over the number of occupied households (Colombia); calculated as the number of interviewed households over the number of targeted households (Peru 2007-8) or over the number of targeted dwellings multiplied by the average number of households per dwellings (Peru 2012-3).

† Anthropometry response rate: calculated as the number of adults with anthropometric measurements (either weight or height) over the number of adults eligible for the measurement (Colombia and Peru); or over the number of adults responding to the questionnaires (Mexico).

‡ $\text{Proportion not responding due to absence} = \dfrac{\text{No. persons not measured due to absence}}{\text{Total no. of eligible persons}}$

among adults aged 20 years or over. The rates were lower among men (61-81%) than women (80-89%). The main reason for no-response was absence at home in both men and women, but its rates among men were roughly double of those among women. This sex difference was observed probably because men were likely to be at work when survey teams visited their residence.

2.2.2 Inclusion and exclusion criteria, and sample sizes

All adults aged 20 to 69 years were included in the study. Adults aged 70 years or over were not included due to their small number of observations. For Colombia, only adults aged 20 to 64 years were included since they did not measure adults aged 65 years or over for anthropometry. Women who were pregnant at the interview or whose records did not have information on pregnancy status were excluded from analyses. Records with height < 110 cm or > 210 cm, weight < 20 kg, body mass index (BMI) < 12 kg/m^2 or > 120 kg/m^2 were excluded as having implausible values possibly due to measurement, recording, or data entry error. Those with either weight or height measurements were excluded since BMI could not be calculated. Those records with implausible or incomplete data were less than 0.1% of all the records except those in the first survey data in Mexico (0.8%) and Peru (0.3-0.4%) (Table 2-2). Final sample sizes were 64,413 for Mexico, 144,628 for Colombia, and 36,082 for Peru.

2.2.3 Response variable

Body mass index (BMI) was used as an indicator to measure overweight and obesity, and its distributions were modeled and assessed. BMI was assumed to follow a Box-Cox power exponential

Table 2-2 Results of exclusions and final sample sizes (adults aged 20-69 years*)

Mexico	2006			2012		
	Men	Women	Total	Men	Women	Total
Measured	12,080	18,740	30,820	14,381	20,454	34,835
Weight or height missing	93	141	234	11	20	31
Values implausible	5	6	11	5	7	12
Pregnant or missing info		416	416		538	538
Eligible for study	11,982	18,178	30,160	14,365	19,888	34,253
Measurement/recording error rate†	0.8%	0.8%	0.8%	0.1%	0.1%	0.1%
Colombia	2005			2010		
	Men	Women	Total	Men	Women	Total
Measured	23,285	37,531	60,816	36,745	49,947	86,692
Weight or height missing	0	0	0	18	23	41
Values implausible	10	22	32	9	8	17
Pregnant or missing info		1,262	1,262		1,530	1,530
Eligible for study	23,275	36,247	59,522	36,718	48,388	85,106
Measurement/recording error rate†	0.0%	0.1%	0.1%	0.1%	0.1%	0.1%
Peru	2007-8			2012-3		
	Men	Women	Total	Men	Women	Total
Measured	8,220	9,823	18,043	8,166	10,465	18,631
Weight or height missing	0	1	1	0	0	0
Values implausible	29	32	61	8	9	17
Pregnant or missing info		284	284		230	230
Eligible for study	8,191	9,507	17,698	8,158	10,226	18,384
Measurement/recording error rate†	0.4%	0.3%	0.3%	0.1%	0.1%	0.1%

* In case of Colombia, adults aged 20-64 years.

† $\text{Measurement/recording error rate} = \dfrac{\text{Number of persons with incomplete or implausible weight, height, or BMI}}{\text{Number of persons with weight and/or height measurement}}$

(BCPE) distribution, whose shape is determined by 4 parameters. The BCPE distribution was developed to provide a model for a response variable that exhibits skewness and kurtosis (Rigby and Stasinopoulos 2004). The BCPE distribution fitted well with Dutch adult BMI distributions (Rigby and Stasinopoulos 2004, Majer et al. 2013) and used for the construction of international child growth curves for weight, length/height, and BMI (WHO 2006).

The BCPE distribution consists of 4 parameters, μ, σ, ν, and τ, which can be interpreted as relating to location (median), scale (approximate coefficient of variation), skewness (transformation to symmetry), and kurtosis (power exponential parameter), respectively (Rigby and Stasinopoulos 2004). A positive random variable Y that follows a BCPE distribution is defined through the transformed random variable Z given by

$$Z = \begin{cases} \dfrac{1}{\sigma\nu}\left[\left(\dfrac{Y}{\mu}\right)^{\nu} - 1\right] & \text{if } \nu \neq 0 \\ \dfrac{1}{\sigma}\log\left(\dfrac{Y}{\mu}\right) & \text{if } \nu = 0 \end{cases}$$

for $0 < Y < \infty$, $\mu > 0$, and $\sigma > 0$, and where the random variable Z is assumed to follow a standard power exponential distribution with a continuous power parameter, $\tau > 0$. The probability density function of Z is given by

$$f_{Z(z)} = \frac{\tau}{c2^{(1+1/\tau)}\Gamma(1/\tau)}\exp(-0.5|z/c|^{\tau})$$

for $-\infty < z < \infty$ and $\tau > 0$, and where $c^2 = 2^{-2/\tau}\Gamma(1/\tau)[\Gamma(3/\tau)]^{-1}$. Then, the probability density function of Y is given by

$$f_Y(y) = \left|\frac{dz}{dy}\right| f_Z(z) = \frac{y^{\nu-1}}{\mu^{\nu}\sigma} f_Z(z)$$

2.2.4 Covariates

Age was a covariate of interest. It was included in analyses as a nominal variable (in 5-year age groups, i.e., 20-24, 25-30, …, 65-69). Time and its interaction with age were included in the model as well. The interaction terms enabled us to estimate age-specific annual change in parameters of the BCPE distribution.

2.2.5 Data preparation

Sampling weights were calibrated after the exclusion of records. First, the sums of sampling weights were calculated by sex and age. Then, after excluding records, sampling weights were calibrated so that the sex- and age-specific sums of sampling weights are equal to those of the original dataset.

Sampling weights were re-calibrated before pooling datasets of two time points in order to accommodate changes in total populations over time and population sizes possibly estimated using different estimation methods when sampling weights were prepared. First, time of each survey was calculated as the mean of anthropometry measurement dates; if the measurement dates were not available, the midpoint of data collection period was used. Second, sex-specific "standard" adult populations aged 20-69 (or 20-64 in case of Colombia) were obtained for two years (as of mid-year, July 1) that surrounded the survey time from the UN population estimates (UN 2015a, 2015b). Then, the standard population size at the survey time was interpolated using the formula for the mean annualized growth rate

$$\bar{r}[T_1, T_2] = \frac{\ln\left[\frac{N(T_2)}{N(T_1)}\right]}{T_2 - T_1}$$

where $N(T_1)$ and $N(T_2)$ are the sizes of population at time T_1 and T_2, respectively (Preston, et al. 2001). Third, sampling weights were re-calibrated so that the sum of sampling weights is equal to the standard population.

After finalizing the calibration of sampling weights, data of two time-points were concatenated separately for county and sex and used for further analyses.

2.2.6 Analysis methods

As an exploratory analysis, histograms of BMI were constructed for each combination of county, survey, sex, and age category. Then, the BCPE distribution, as well as the log-normal and the normal distributions as a comparison, were fitted to the data of each survey by sex and age. The estimated distribution curves were superimposed on the histograms of observed BMI to verify the goodness-of-fit of the BCPE distributions to the data. The fitted model was an intercept-only model, in which each parameter of the BCPE distribution was regressed without covariates, and its model was given by

$$BMI \sim BCPE(\mu, \sigma, \nu, \tau)$$

$$\mu = \beta_0^{\mu}, \qquad \log\sigma = \beta_0^{\sigma}, \qquad \nu = \beta_0^{\nu}, \qquad \log\tau = \beta_0^{\tau}$$

The above model was fitted using a regression model called the generalized additive model for location, scale and shape (GAMLSS). With this model, not only the location (mean or median) but also other parameters of the distribution of the response variable can be modeled as parametric functions of covariates. Maximum likelihood estimation is used to fit the model (Rigby and Stasinopoulos 2005).

In order to assess magnitude and rates of change in BMI distributions, the GAMLSS regression model was applied. To compare the magnitude across the countries, BMI distributions were estimated by age group for the year 2010 separately by country and sex. To assess the rates of change in BMI distributions, the distributions in the years 2005 were estimated and compared with those in the year 2010. The fitted model was

$$BMI \sim BCPE(\mu, \sigma, \nu, \tau)$$

$$\mu = \sum\nolimits_k \left[{}^{\mu}\beta_{1k} age_k + {}^{\mu}\beta_{2k}(age_k \times time) \right] = \boldsymbol{X}_1 \, {}^{\mu}\boldsymbol{\beta}_1 + \boldsymbol{X}_2 \, {}^{\mu}\boldsymbol{\beta}_2$$

$$\log\sigma = \sum\nolimits_k \left[{}^{\sigma}\beta_{1k} age_k + {}^{\sigma}\beta_{2k}(age_k \times time) \right] = \boldsymbol{X}_1 \, {}^{\sigma}\boldsymbol{\beta}_1 + \boldsymbol{X}_2 \, {}^{\sigma}\boldsymbol{\beta}_2$$

$$\nu = \sum\nolimits_k \left[{}^{\nu}\beta_{1k} age_k + {}^{\nu}\beta_{2k}(age_k \times time) \right] = \boldsymbol{X}_1 \, {}^{\nu}\boldsymbol{\beta}_1 + \boldsymbol{X}_2 \, {}^{\nu}\boldsymbol{\beta}_2$$

$$\log\tau = \sum\nolimits_k \left[{}^{\tau}\beta_{1k} age_k + {}^{\tau}\beta_{2k}(age_k \times time) \right] = \boldsymbol{X}_1 \, {}^{\tau}\boldsymbol{\beta}_1 + \boldsymbol{X}_2 \, {}^{\tau}\boldsymbol{\beta}_2$$

where

age_k: the indicator variable for the age category k

$time$: the continuous variable for time centered at July 1, 2010

${}^{i}\beta_{1k}$: the coefficient for the age category k for the parameter $i = (\mu, \sigma, \nu, \tau)$

${}^{i}\beta_{2k}$: the coefficient for the interaction between the age category k and time for the parameter i

$\boldsymbol{X}_1$: the vector of covariates for the age group

$\boldsymbol{X}_2$: the vector of covariates for the interaction between the age group and time

${}^{i}\boldsymbol{\beta}_1$: the vector of coefficients for the age group for the parameter i

${}^{i}\boldsymbol{\beta}_2$: the vector of coefficients for the interaction between the age group and time for the parameter i

Time was centered at July 1, 2010 so that coefficient estimates for the age terms represent the values as of mid-year 2010. The interaction terms represent the annual rates of change in the BCPE parameters. After fitting the model separately for country and sex, values of 4 BCPE parameters were estimated for each combination of year (2005 or 2010) and age group (20-24, 25-29, …, or 65-69). And then, estimated BMI distribution curves (i.e., density cures) were constructed. In order to quantify the estimated BMI distributions, prevalence of 4 BMI categories was calculated from the estimated cumulative density function $\hat{F}_{BMI}$. The 4 BMI categories were: undernutrition (BMI < 18.5 kg/m^2); normal (18.5 ≤ BMI < 25.0 kg/m^2); overweight (25.0 ≤ BMI < 30.0 kg/m^2); and obese (BMI ≥ 30 kg/m^2), which can be given by $\hat{F}_{BMI}(18.5)$, $\left[\hat{F}_{BMI}(25.0) - \hat{F}_{BMI}(18.5)\right]$, $\left[\hat{F}_{BMI}(25.0) - \hat{F}_{BMI}(18.5)\right]$, and $\left[\hat{F}_{BMI}(30.0) - \hat{F}_{BMI}(25.0)\right]$, respectively (Majer et al. 2013).

Sampling weights and clustering within primary sampling units (PSUs) were incorporated in all analyses. In order to estimate variances accounting for clustering at the sampling unit level, 2,000 bootstrap samples were generated and used. Bootstrap technique was applied since the GAMLSS package did not allow models with clustering. A bootstrap sample was obtained by sampling $(n_h - 1)$ PSUs with replacement per stratum for all strata, where n_h stands for the number of PSUs in the stratum h (Kolenikov 2010). A stratum was defined as a combination of the sub-administrative unit (states in Mexico and departments in Colombia and Peru) and the urban-rural area category (for Mexico, the metropolitan-urban-rural area category). Definitions of urban and rural areas differed by countries (Appendix C), but these country-specific definitions were used in analyses. In the case of Colombia, secondary sampling units (SSUs) were used instead of PUSs for bootstrap sampling purposes to incorporate its peculiar sampling design (Appendix C). For each bootstrapped sample, sampling weights were calibrated so that sex- and age-specific sums of sampling weights are equal to those of the original data.

Let $\widehat{\boldsymbol{\theta}}_{(i)}$ be the vector of coefficient estimates from the i^{th} replication, and the total number of replications be r. Then, the variance-covariance was estimated by

$$\widehat{\boldsymbol{V}}(\widehat{\boldsymbol{\theta}}) = \frac{1}{r}\sum_{i=1}^{r}\{\widehat{\boldsymbol{\theta}}_{(i)} - \overline{\boldsymbol{\theta}}_{(.)}\}\{\widehat{\boldsymbol{\theta}}_{(i)} - \overline{\boldsymbol{\theta}}_{(.)}\}'$$

where $\bar{\theta}_{(.)}$ is the bootstrap mean

$$\overline{\boldsymbol{\theta}}_{(.)} = \frac{1}{r}\sum_{i=1}^{r}\widehat{\boldsymbol{\theta}}_{(i)}$$

(Rao et al. 1992). This estimated variance-covariance was used for the hypothesis testing. Confidence intervals were reported as 95% bootstrap percentile confidence intervals. The model did not converge in 100 iterations with some bootstrap samples (11 (0.5%) and 29 (1.5%) out of 2,000 samples from the Mexican male and female data, respectively; and 1 (0.1%) from the Peruvian male data). For these samples, an average of estimates from the last 50 iterations was used as the final estimate after verifying that the estimates of each iteration were not diverging.

R (version 3.2.2) was used to fit the GAMLSS regression models, and Stata (version 14.1) was used for the rest of the analyses.

2.3 Results

2.3.1 Study population

No notable differences were observed in age composition between two surveys except Peru where the proportion of the population aged 50 years or over increased in the second survey as compared to the first survey (Table 2-3). About 75% or more adults aged 20-69 years lived in urban areas in the second survey in all the 3 countries. No notable differences were observed in the urban-rural ratio except Peru where the urban population increased by 7 percentage points between two surveys. Across the 3 countries, the proportion of people with higher educational attainment increased whereas the proportion of people with primary education or less declined between 2 surveys.

Table 2-3 Demographic characteristics of the study populations (adult aged 20-69 years*)†

	Mexico (n = 64,413)				Colombia (n = 144,628)				Peru (n = 36,082)			
	Men (n = 26,347)		Women (n = 38,066)		Men (n = 59,993)		Women (n = 84,635)		Men (n = 16,349)		Women (n = 19,733)	
	2006	2012	2006	2012	2005	2010	2005	2010	2007-8	2012-3	2007-8	2012-3
Number of observations												
Number of records	11,982	14,365	18,178	19,888	23,275	36,718	36,247	48,388	8,191	8,158	9,507	10,226
Weighted counts (000)	29,166	33,855	29,680	33,980	11,763	13,096	11,973	13,382	7,773	8,304	7,607	8,067
Age (%)												
20-29	29	29	27	26	33	30	30	28	29	24	27	22
30-39	25	24	27	27	25	25	26	24	25	22	27	23
40-49	21	21	22	21	22	23	23	23	24	23	24	24
50-59	15	16	14	16	15	16	16	18	13	18	14	19
60-69	10	10	10	10	5	6	6	6	9	13	9	13
Type of residence (%)												
Urban	78	79	77	79	73	74	77	78	68	75	71	78
Rural	22	22	23	21	27	26	23	22	32	25	29	22
Educational attainment (%)												
Primary	44	34	53	40	41	38	42	36	28	24	39	35
Secondary	41	48	37	45	40	42	41	42	43	41	35	35
Higher	14	18	10	16	17	20	18	22	29	34	26	30
Missing	1	0	0	0	1	1	0	0	0	0	0	0
Household wealth quartile (%)												
Lowest	22	23	24	23	25	26	21	21	21	23	20	21
Lower	25	24	25	25	26	25	26	25	24	24	23	23
Higher	26	25	26	26	25	25	27	27	26	23	27	25
Highest	27	27	25	26	25	24	27	27	25	25	27	25
Missing	0	0	0	0	0	0	0	0	4	6	4	7

* In case of Colombia, adults aged 20-64 years.

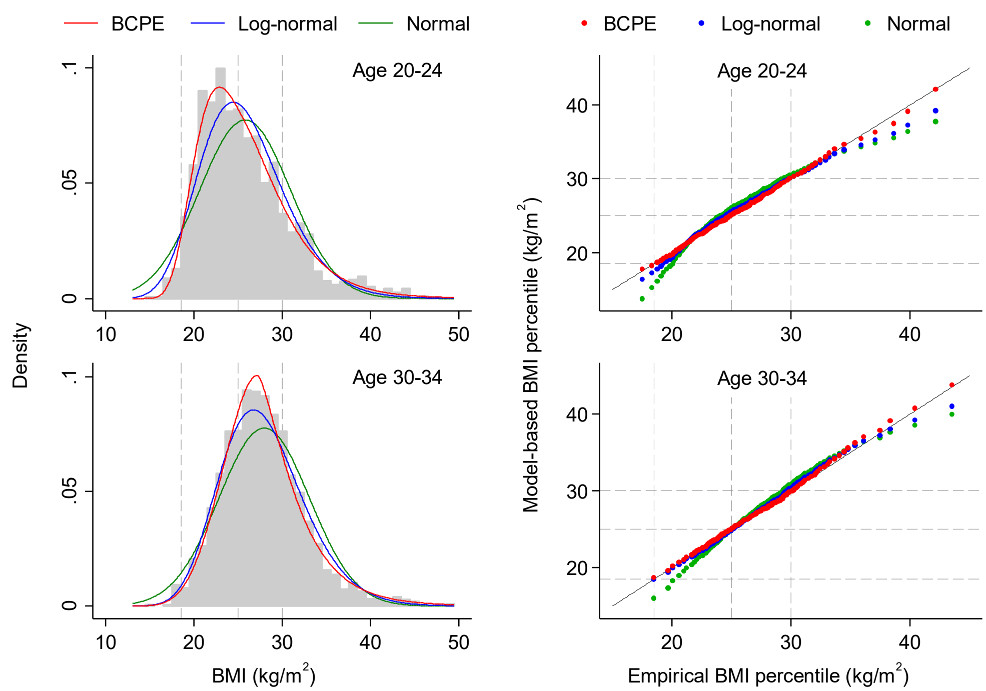
† All the numbers were calculated using sampling weights except the number of records.

2.3.2 Fit of the BCPE distribution to BMI data

The BCPE distribution fitted well on the BMI data across countries, sex, and age. The example shown in Figure 2-1 is from the BMI data of Mexican men measured in the 2012 survey. These right-skewed BMI distributions were a commonly observed pattern among the young populations across the 3 countries. BMI density curves estimated with the BCPE distribution overlapped well with the observed BMI distributions as compared to those estimated with the log-normal and the normal distributions. The BCPE distribution fitted better, especially at the lower and higher ends of the distributions.

Figure 2-1 Observed BMI distributions and estimated BMI density curves with the BCPE, log-normal, and normal distributions

(Example from Mexican men aged 20-24 and 30-34 in the 2012 survey)



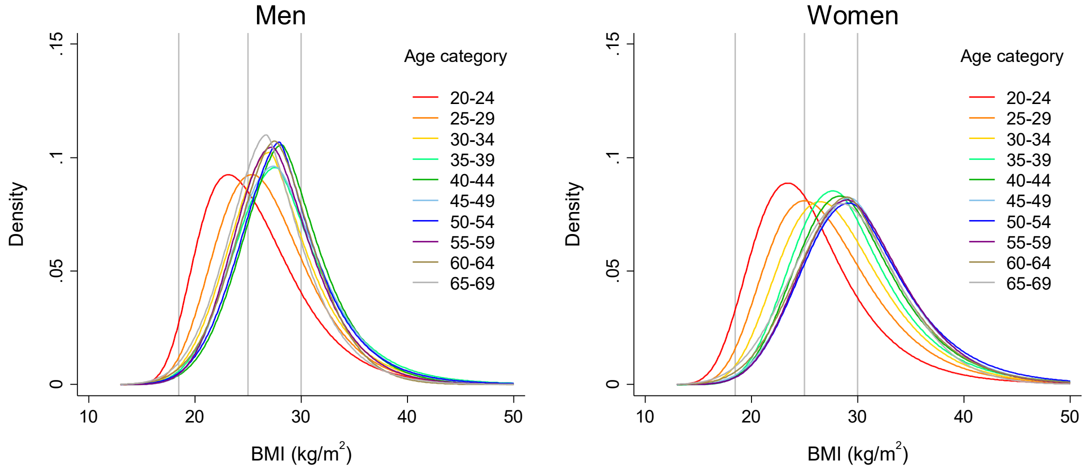
* Vertical dashed lines show the cutoffs for BIM categories of underweight, normal, overweight, and obesity.

2.3.3 Estimated BMI distributions in 2010

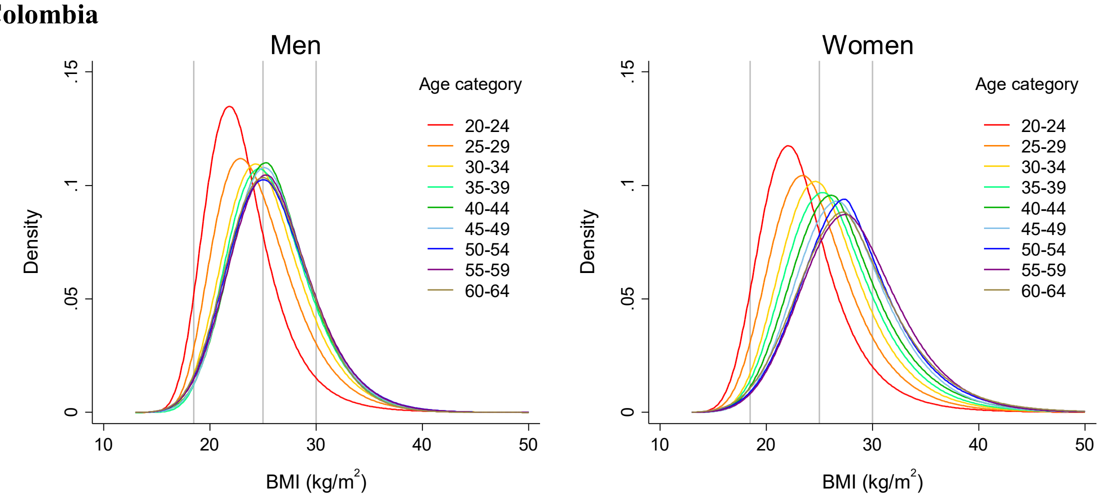
Estimated BMI distributions and prevalence of 4 BMI categories for the year 2010 are presented in Figure 2-2 and 2-3, and their difference between the years 2005 and 2010, in Figure 2-4 and 2-5. Estimated parameters are found in Appendix D. Across the 3 countries in 2010, BMI was distributed wider and located at higher levels at all ages in women as compared to men. BMI distributions continued shifting upward and becoming wider by age, and this trend continued until about age mid- to late-30's in men and about mid- to late-40's in women (Figure 2-2). BMI was distributed similarly in men and women aged 20-24 years in Mexico whereas women's BMI was higher in Colombia and Peru, but noting that there were already more young adults with high BMI in Mexico as compared to the other 2 countries. Overweight and obesity prevalence was at the highest around age 40's in men and around 50's in women (Figure 2-3).

Figure 2-2 Estimated BMI distributions by age for year 2010

Mexico



Colombia



Peru

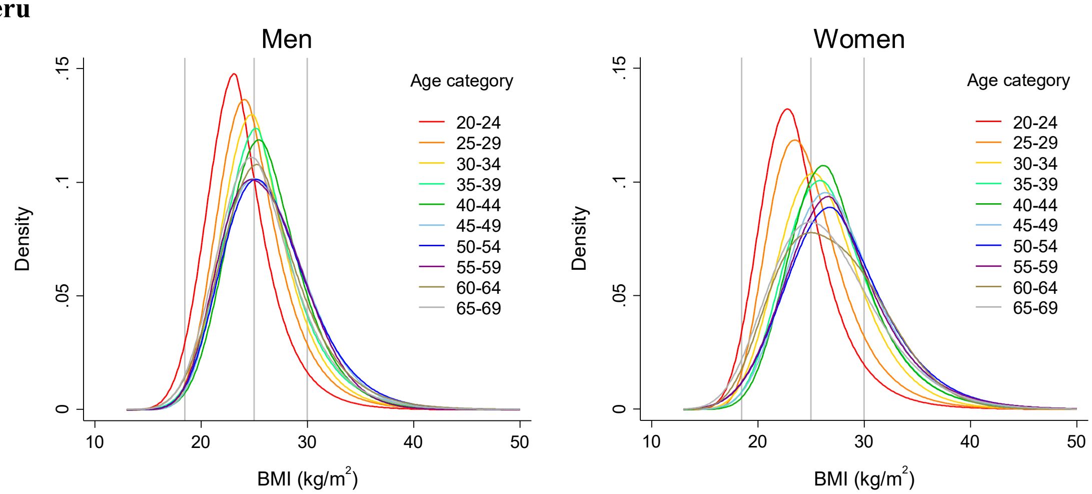
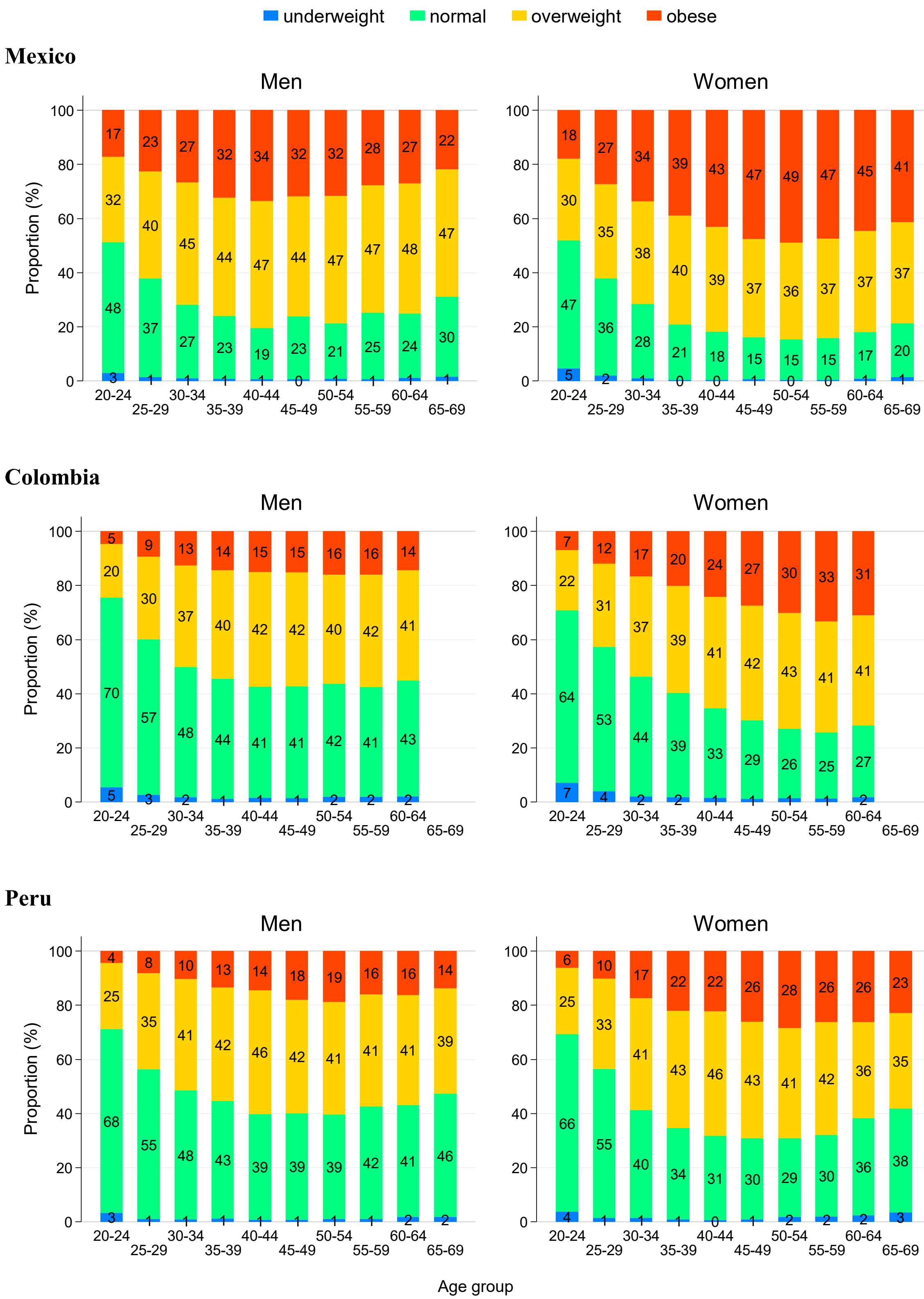


Figure 2-3 Estimated prevalence of 4 BMI categories by age for year 2010



In Mexico, in men, the lower end of the BMI distribution curve moved upward by age until the early 30's whereas the upper end did not move much (Figure 2-2, top left). By contrast, in women, the entire distribution shifted upward by age until early 40's (Figure 2-2, top right). In other words, adult men with normal or lower BMI increased their BMI relative to others until early 30's whereas all women tended to increase their BMI until age 40's, if there is little cohort effect. The highest overweight and obesity prevalence was among age group 40-44 yeas in men (80.5%, 95% percentile interval: 78.5-82.4) and among age group 50-54 years in women (84.7%, 83.0-86.3) (Figure 2-3, top).

In Colombia, in both men and women, the BMI distributions shifted upward and flattened by age. This pattern stopped around late 30's in men whereas it continued until early 50's in women (Figure 2-2, middle). The highest overweight and obesity prevalence was among age group 40-44 years in men (57.3%, 55.6-59.1) and among age group 55-59 years in women (74.3%, 72.7-76.0) (Figure 2-3, middle).

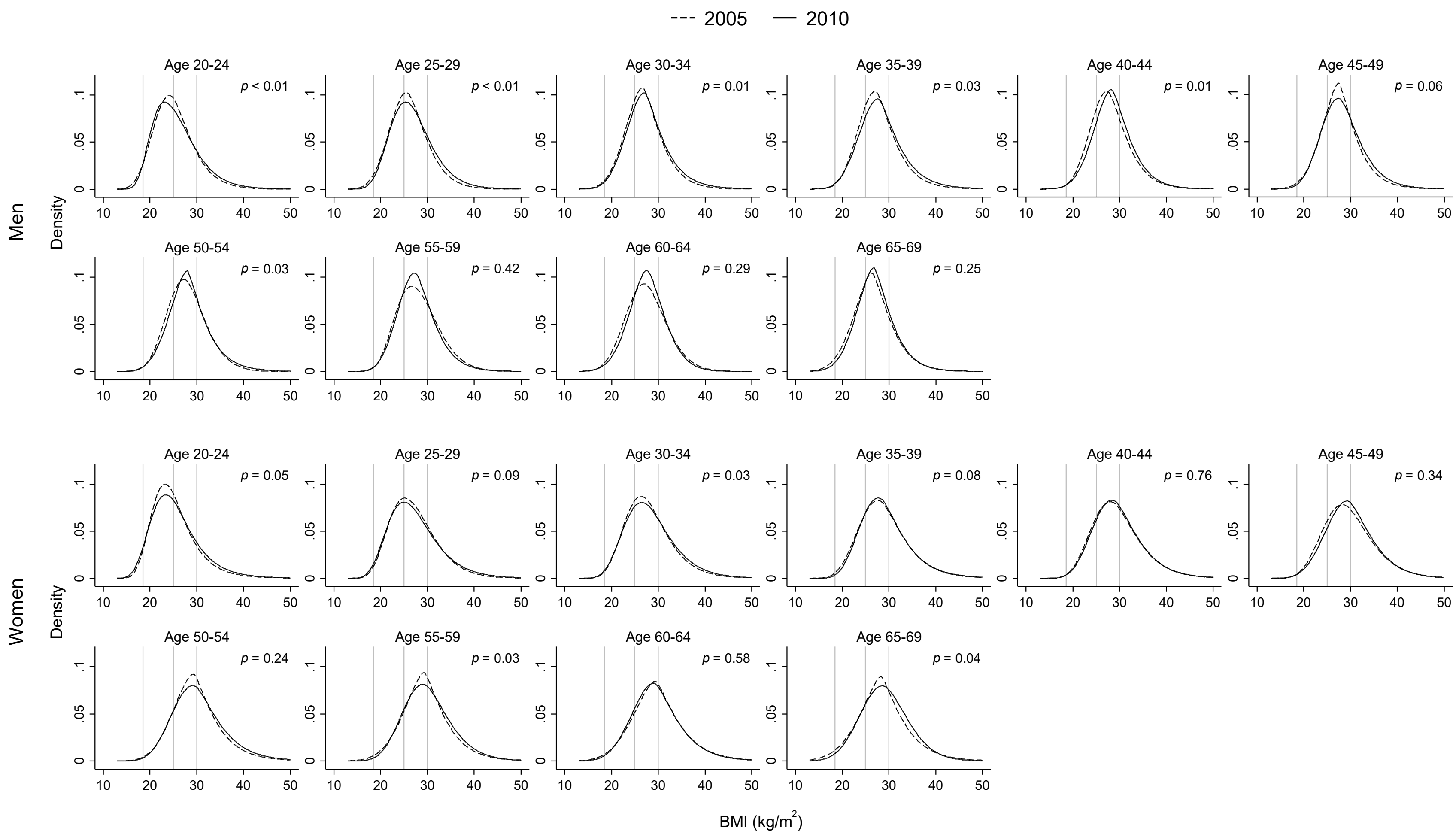
In Peru, in both men and women, the BMI distributions shifted upward and flattened by age, similar to Colombia but more sharply. The change in the BMI distributions stopped around mid-40's in both men and women (Figure 2-2, bottom). The highest overweight and obesity prevalence was among age group 50-54 years in both men and women, 60.4% (60.4%, 57.5-63.2) and 69.1% (66.4, 71.6), respectively (Figure 2-3, bottom).

2.3.4 Estimated differences in BMI distributions between 2005 and 2010

Across the 3 countries, young- to middle-aged adults aged 25-49 years increased significantly their BMI between year 2005 and 2010 except Mexican women (Figure 2-4, -5). In Colombia and Peru, significant increases were observed among the elderly aged 55 years and over as well. Men generally had faster increases in BMI than women, more apparently in Mexico and Peru (Figure 2-5).

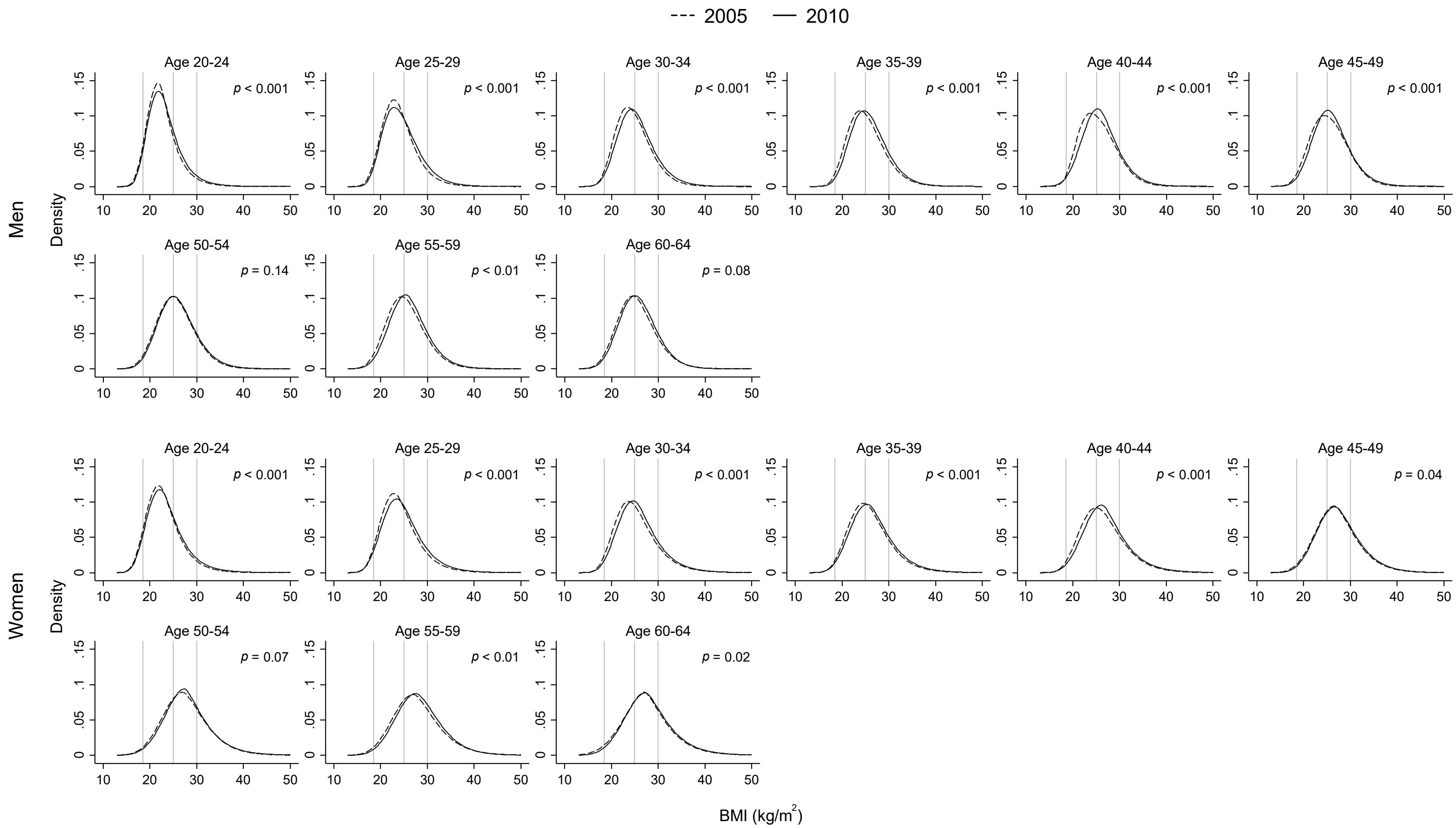
The most notable differences in BMI distributions was observed in Peru (Figure 2-4c). In men, the distributions rather shifted upward in almost all age groups, in other words, men generally increased their BMI between 2005 and 2010 on average. In women, increases in proportion of higher end of the BMI distribution were observed below age 30 years, and shifts of the entire distributions were observed

Figure 2-4a Estimated BMI distributions by age for years 2005 and 2010, Mexico



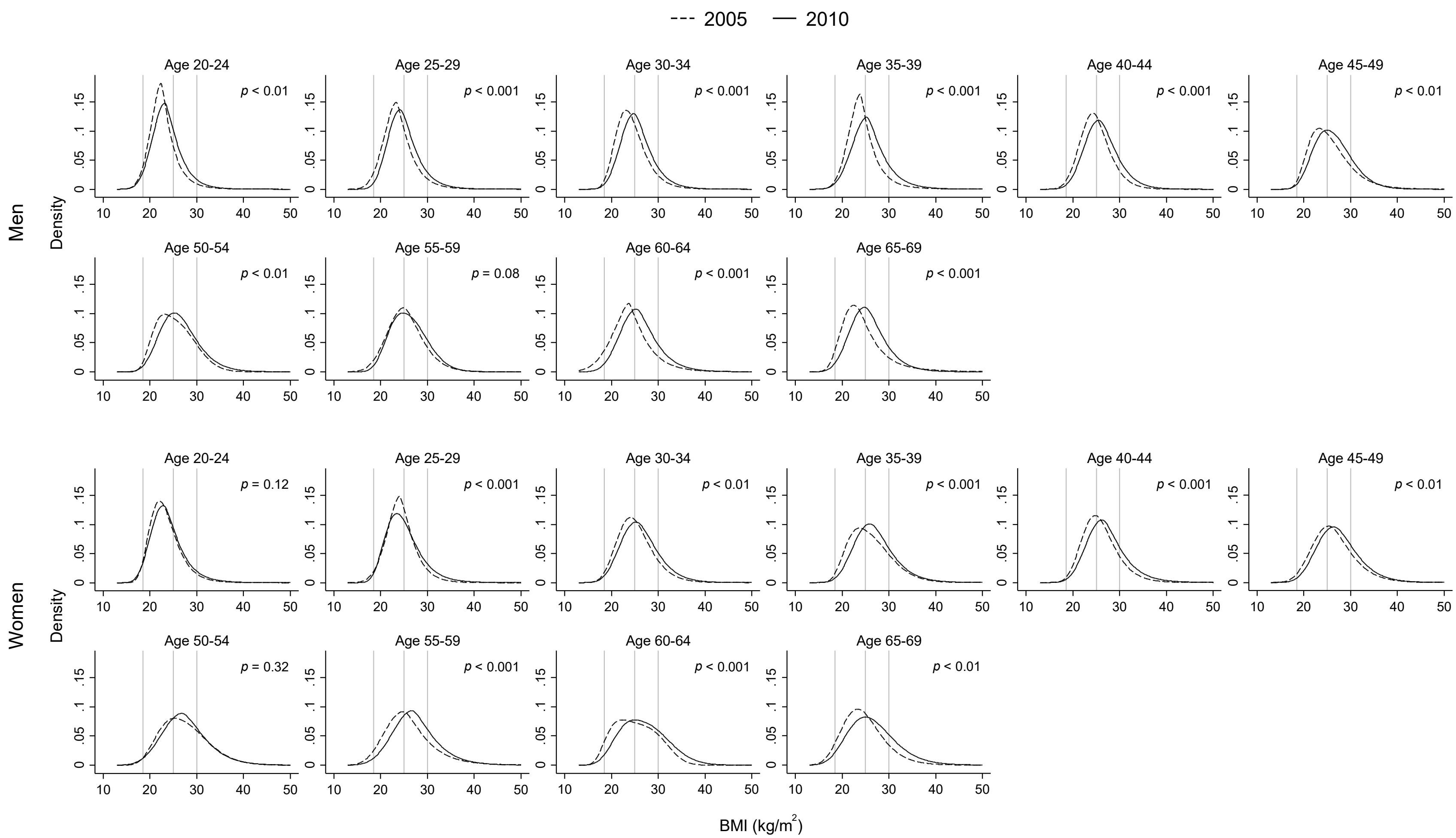
* P-values (from 4 *df* likelihood ratio tests) indicate difference in distributions between years 2005 and 2010. The null hypothesis is: $\beta_{2j}^{\mu} = 0, \beta_{2j}^{\sigma} = 0, \beta_{2j}^{\nu} = 0, \beta_{2j}^{\tau} = 0$ where β_{2j}^{i} is the coefficient for the interaction term between age group j and time for the parameter $i = (\mu, \sigma, \nu, \tau)$.

Figure 2-4b Estimated BMI distributions by age for years 2005 and 2010, Colombia



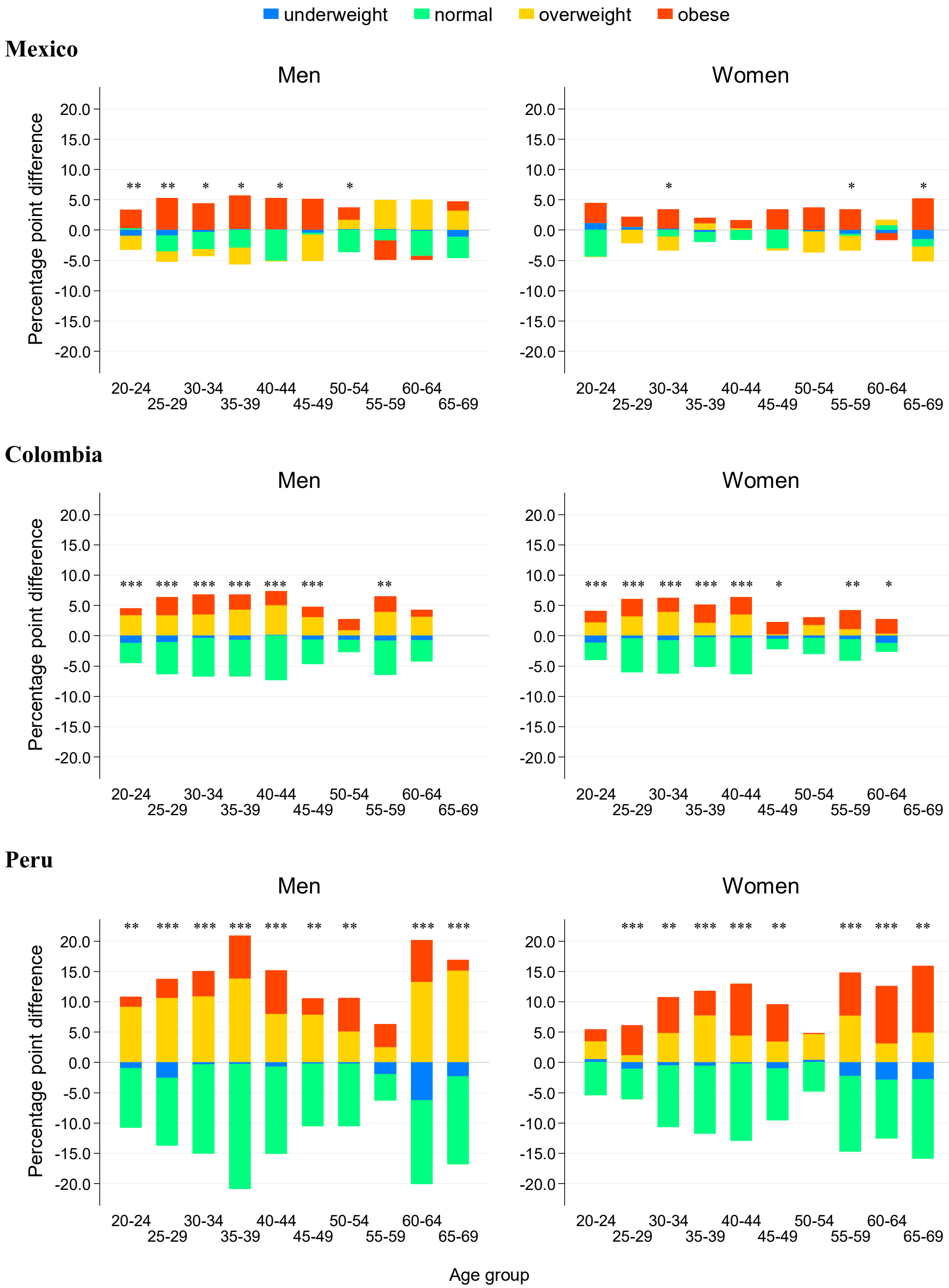
* *P*-values (from 4 *df* likelihood ratio tests) indicate difference in distributions between years 2005 and 2010. The null hypothesis is: $\beta_{2j}^{\mu} = 0, \beta_{2j}^{\sigma} = 0, \beta_{2j}^{\nu} = 0, \beta_{2j}^{\tau} = 0$ where β_{2j}^{i} is the coefficient for the interaction term between age group j and time for the parameter $i = (\mu, \sigma, \nu, \tau)$.

Figure 2-4c Estimated BMI distributions by age for years 2005 and 2010, Peru



* P-values (from 4 df likelihood ratio tests) indicate difference in distributions between years 2005 and 2010. The null hypothesis is: $\beta_{2j}^{\mu} = 0, \beta_{2j}^{\sigma} = 0, \beta_{2j}^{\nu} = 0, \beta_{2j}^{\tau} = 0$ where β_{2j}^{i} is the coefficient for the interaction term between age group j and time for the parameter $i = (\mu, \sigma, \nu, \tau)$.

Figure 2-5 Change in prevalence of 4 BMI categories by age between 2005 and 2010



* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$ in the test for the difference in distributions between 2005 and 2010 from 4 *df* likelihood ratio tests. The null hypothesis is: $\beta_{2j}^{\mu} = 0, \beta_{2j}^{\sigma} = 0, \beta_{2j}^{\nu} = 0, \beta_{2j}^{\tau} = 0$ where β_{2j}^{i} is the coefficient for the interaction term between age group j and time for the parameter $i = (\mu, \sigma, \nu, \tau)$. See Appendix D for actual *p*-values.

about age 30 years and over. Among middle-age people, the highest percentage point (pp) increase in overweight and overweight prevalence was 20.9 pp (12.2, 30.1) among men aged 35-39 years and 13.0 pp (7.3, 18.8) among women aged 40-44 years (Figure 2-5).

2.4 Discussion

2.4.1 Fit of the BCPE distribution to BMI data

The BCPE distributions provided a good fit for the BMI data of the 3 countries, and the assumption that BMI follows a BCPE distribution seemed reasonable. This assured that we are able to discuss further the following results about estimated BMI distributions. On the other hand, the log-normal distribution did not fit well at the lower and higher ends of the BMI distribution, which were similar results to those with the Dutch data (Majer et al. 2013). This seemed an advantage of a flexible model that incorporates not only the location and scale of the distribution but also its shape (skewness and kurtosis) as well.

2.4.2 Main findings

BMI distributions were wider among women than men in the 3 countries. This result overlapped with a previous report that mean BMI varied more among women than men in 2014 across countries in the world (NCD Risk Factor Collaboration 2016). Among the 3 countries assessed in our study, women were generally with higher BMI at almost all ages, and more notably in the middle-age or above. This sex-difference was also consistent with previous findings in which mean BMI and obesity prevalence were higher in women than men (Schargrodsky et al. 2008, Miranda et al. 2013, NCD Risk Factor Collaboration 2016). However, it should be noted that BMI is not necessarily higher in women than men across countries, and the degree of sex differences vary, as observed in the introduction (1.1 Trends of adult overweight and obesity in low- and middle-income countries and in Latin America).

Whereas overweight and obesity prevalence was lower in Peru than the other 2 countries, its rate of increase was much higher especially among middle-aged adults (about age 30-35) and the elderly in men and women. Researchers in Peru list possible causes of wide spreading overweight and obesity, which include improved economy, sociocultural change, availability of cheap food rich in carbohydrate, fat, and sugar, existence of poverty and food access limited to such cheap but energy-dense food (Lanata 2012, Álvarez-Dongo et al. 2012).

The rate of increase in overweight and obesity was higher in men than women, more notably in Peru and Mexico. This observed trends had a similarity with the report from Brazil in which change in prevalence between 1975, 1989, and 2003 was assessed (Monteiro et al. 2007). The authors demonstrated that a steady increase in obesity prevalence was observed in 2 consecutive inter-survey periods in men whereas an increase was observed only between 1975 and 1989 in women. The halt in rise among women was due to a decrease in obesity prevalence in high-income groups.

2.4.3 Implications for public health practice

Whereas an assessment of BMI distributions at a specific time identifies subpopulations that are currently most affected, an assessment of rates of change in BMI distributions identifies those who become more affected. When health policies and programs are planned, both aspects should be considered, and the populations might be targeted differently. Subpopulations that are already affected require more curative or "secondary prevention" against the development of other obesity-associated risk factors and chronic diseases. Examples may include weight reduction programs and continuous care at health facilities for control of weight and co-existing problems. Subpopulations that become more affected over time may require preventive strategies to minimize the increase of overweight and obesity, which may include health education and communication regarding diet and exercise for weight control, and screening and consultation at health facilities.

Differences in BMI distributions and their changes by sex and age imply necessity of tailoring health policies and programs if and where appropriate. Among the 3 countries assessed in our study, men generally increase their BMI at their earlier stage of adulthood about age 20-35 whereas women tended to increase their BMI in a prolonged manner until age 40-50. It was also noticed that rates of change in BMI distributions were higher in men than women and in young- to middle-aged people. Besides, BMI distributions of the youngest age group 20-24 differed among the 3 countries and/or between two sexes, implying that some of them already had faced a problem of overweight and obesity at their early age. Hence, it would be necessary to think which sex and age groups should be targeted and which aspect of the problem should be tackled (i.e., preventive vs. curative), which in turn would determine types of interventions to be implemented.

High values of BMI among middle-age populations and increases in BMI among young and middle-age adults could create huge challenges to the health care systems of these countries. Increases of overweight and obese populations among young and middle-age adults imply potential further increases of obesity-related diseases, such as cardiovascular diseases and diabetes, which will be even accelerated by prolonged life expectancy. Since such diseases are often accompanied with chronic conditions that require continuous and expensive control and treatment, financial burdens for the individuals and the health systems could increase enormously. In addition, these chronic diseases progress gradually and cannot be cured instantly once they are diagnosed, early mitigation of overweight and obesity problems and forward-thinking preparation for future burdens of such diseases on populations and health systems would be indispensable.

2.4.4 Strengths and limitations

There are some strengths of our study. The study was based on the data of nationally representative surveys that included both men and women and assessed countries in a comparative manner by fixing the same time period. Also, the study assessed magnitude and rates of change in

BMI distributions to identify age groups that suffer from overweight and obesity currently as well as those being more affected than previously. The entire BMI distributions were modeled in place of obesity indicators that represent partial aspect of the distribution. Visualized BMI distributions could help better understanding of dynamicity of the obesity distributions. On the analytical side, from our best knowledge, our study is the first attempt to apply the BCPE-GAMLSS technique for survey data considering both sampling weights and clustering. Hence, this technique should be applicable to other countries that have population survey data with anthropometric measurement.

There are some limitations of our study as well. First, the findings are based on two cross-sectional studies, and therefore, the association patterns we observed between age and BMI distributions consist of age effects and cohort effects. For example, it is usually not true that those in age group 20-24 will experience the same distributions as observed at higher age groups. In fact, a report from the US indicated that the younger generations were with higher BMI as compared to older generations of the same age (Wang et al. 2007). Also, there might be ongoing interventions that affected nutritional status of the populations unevenly across ages. For example, a conditional cash transfer program for the improvement of maternal and child health conditions might facilitate an increase in weight of family members of childbearing age who are eligible for the programs. Second, differences in BMI distributions were evaluated by comparing the estimated distributions between 2005 and 2010. The 5-year interval may be too short to assess the differences although there were large and significant changes observed in Peru. It would be a next task to assess such changes at longer intervals when more data become available. Third, the interpretation might be handled with caution, especially for men, since their response rates were low in some surveys (e.g., 61% among men in Colombian survey in 2005). Low response rate might be due to the fact that men tend to be absent from home due to work in general. It could be possible that those who worked outside home might be different from those stayed at home during the survey visit time (although survey teams tried several timeslots of different days).

Our study presented differences in magnitude and rates of change in BMI distributions by sex and age. Further analyses are needed at country level in order to identify determinants of observed differences.

Chapter 3 Magnitude and rates of change in overweight and obesity distributions by geographic and socioeconomic factors

3.1 Objective

The objective of this component of the study is to assess differences in magnitude and rates of change in BMI distributions by geographic and socioeconomic factors in both sexes.

3.2 Methods

3.2.1 Data sources

Data sources for this study were the same as those used in the previous chapter. They were the data of two time points from nationally representative health and nutritional household surveys conducted between 2005 - 2013 in Mexico, Colombia, and Peru. Detailed description of each survey was given in the previous chapter (2.2.1 Data sources).

3.2.2 Inclusion and exclusion criteria, and sample sizes

All adults aged 20 to 59 years were included in the study. The study population was limited to the adult population aged 20-59, instead of 20-69, to be able to compare results across countries (Colombia had data for populations aged 20-64), to reduce convergence problem due to small sample size in the population aged 60-69, and to be able to better assume equal effects of geographic and socioeconomic factors on BMI distribution across age groups. Other exclusion criteria were implausible values for weight, height and BMI, and pregnant women, the same as the previous chapter (2.2.2 Inclusion and exclusion criteria, and sample sizes). Some observations did not carry information on educational attainment or household wealth, and such observations excluded due to missing values were 0.0-0.5% in Mexico, 0.3-1.2% in Colombia, and 2.7-3.6% in Peru among those with valid BMI (Table 3-1). Final sample sizes were 57,414 for Mexico, 135,403 for Colombia, and 30,811 for Peru.

Table 3-1 Results of exclusion and final sample sizes (adults aged 20-59 years)

Mexico	2006			2012		
	Men	Women	Total	Men	Women	Total
Observations with valid BMI	10,672	16,444	27,116	12,643	17,809	30,452
Missing residence info	0	0	0	0	0	0
Missing education info	41	43	84	0	0	0
Missing wealth info	17	44	61	5	5	10
Eligible for study	10,614	16,358	26,972	12,638	17,804	30,442
Exclusion rate due to missing*	0.5%	0.5%	0.5%	0.0%	0.0%	0.0%
Colombia	2005			2010		
	Men	Women	Total	Men	Women	Total
Observations with valid BMI	21,922	34,311	56,233	34,474	45,535	80,009
Missing residence info	0	0	0	0	0	0
Missing education info	265	92	357	342	140	482
Missing wealth info	0	0	0	0	0	0
Eligible for study	21,657	34,219	55,876	34,132	45,395	79,527
Exclusion rate due to missing*	1.2%	0.3%	0.6%	1.0%	0.3%	0.6%
Peru	2007-8			2012-3		
	Men	Women	Total	Men	Women	Total
Observations with valid BMI	7,285	8,547	15,832	6,997	8,953	15,950
Missing residence info	0	0	0	0	0	0
Missing education info	4	4	8	1	0	1
Missing wealth info	199	226	425	218	321	539
Eligible for study	7,083	8,318	15,401	6,778	8,632	15,410
Exclusion rate due to missing*	2.8%	2.7%	2.7%	3.1%	3.6%	3.4%

$$* \text{ Exclusion rate due to missing} = \frac{\text{Number of persons with missing values}}{\text{Number of persons with valid BMI}}$$

3.2.3 Response variable

BMI was used as an indicator to measure overweight and obesity as in the previous chapter. BMI was assumed to follow a Box-Cox power exponential (BCPE) distribution.

3.2.4 Covariates

Included covariates in regression analyses were place of residence (urban or rural), educational attainment, household wealth, and 10-year age group. Definitions of urban and rural areas varied by country, and country-specific definitions were used in the analyses (see Appendix C for country-specific definitions). Education and household wealth were selected as socioeconomic indicators since they are commonly used indicators in evaluation of overweight and obesity in low- and middle-income countries. Occupation, another indicator of socioeconomic status, was not used because of difficulty in grouping various occupational categories and in handling different occupations among household members. For the educational attainment, 3 categories were used (primary school or less, secondary school, and higher education, including both complete and incomplete cases). For Mexico, preparatory and basic normal school were categorized as secondary school; and superior normal school was categorized as higher education. For household wealth, a household wealth index was constructed, and its quartiles were used for analyses. Household wealth was selected, instead of income, since household income data were not available in Colombian data and household wealth has been commonly used in previous studies. In addition, it was reported that household income is difficult to measure accurately due to various reasons, which include inaccurate or false reported values, difficulty in obtaining income from all household members, and fluctuation in income (Rutstein and Kiersten 2004). For age, 10-year age groups, instead of 5-year age groups used in the previous chapter, were included for the adjustment purpose. A wider age range was used to reduce the number of parameters to be estimated and since age was not a variable of interest.

A household wealth index was constructed from variables on dwelling characteristics, available household services and assets following an established method by the Demographic Health Surveys (DHS) Program (Rutstein and Kiersten 2004, Rutstein 2008, DHS Program n.d., Fry et al. 2014. See Appendix E for the detailed procedure). The wealth index was constructed for each survey (i.e., relative index), instead of an absolute index that is comparable across surveys within a country or across the countries. We chose the relative index since we aimed to assess changes in prevalence among the same segments of the population at each time, classified by quartile of the index, rather than among subpopulations with the same wealth. The household wealth index/score was calculated with principal component analysis (PCA). The first principal component was taken as the underlying wealth score, and the household wealth score was calculated with the PCA weights, separately for urban and rural households. Urban and rural wealth scores were combined into a national score according to the method described elsewhere (Rutstein 2008). Included variables in PCA were number of people per bedroom, availability of domestic servants, house type, construction materials, available services (water, electricity, sewage system, cooking fuels), household items, equipment and vehicles, and animal stock and land area, if any. Quartiles of the wealth score were used, instead of commonly used quintiles, to reduce the number of variables in the model in comparison to the variable for education that had 3 categories.

3.2.5 Data preparation

Sampling weights were calibrated in the same way as described in the previous chapter (2.2.5 Data preparation). The calibration was done in order to maintain the same sex- and age-specific population sizes before and after the exclusion of records and to accommodate changes in total populations over time and population sizes possibly estimated using different estimation methods in preparing sampling weights for each survey. After the calibration of sampling weights, data of two time-points were concatenated separately for country and sex and used for further analyses.

3.2.6 Analysis methods

In order to assess magnitude and rates of change of BMI distributions, the GAMLSS regression model was applied following a similar method applied in the previous chapter (2.2.6 Analysis methods). To compare the magnitude of overweight and obesity across countries, BMI distributions were estimated by area of residence (urban or rural), educational attainment, and household wealth quartile for each 10-year age group for the year 2010 separately by country and sex. To assess the rates of change in overweight and obesity, BMI distributions in the years 2005 were estimated and compared with those in the year 2010. The fitted model was

$$BMI \sim BCPE(\mu, \sigma, \nu, \tau)$$

$$\mu = \sum_{j=1}^{4} \left[\boldsymbol{X}_{1j}\, {}^{\mu}\boldsymbol{\beta}_{1j} + \boldsymbol{X}_{2j}\, {}^{\mu}\boldsymbol{\beta}_{2j} \right]$$

$$\log \sigma = \sum_{j=1}^{4} \left[\boldsymbol{X}_{1j}\, {}^{\sigma}\boldsymbol{\beta}_{1j} + \boldsymbol{X}_{2j}\, {}^{\sigma}\boldsymbol{\beta}_{2j} \right]$$

$$\nu = \sum_{j=1}^{4} \left[\boldsymbol{X}_{1j} \, {}^{\nu}\boldsymbol{\beta}_{1j} + \boldsymbol{X}_{2j} \, {}^{\nu}\boldsymbol{\beta}_{2j} \right]$$

$$\log \tau = \sum_{j=1}^{4} \left[\boldsymbol{X}_{1j} \, {}^{\tau}\boldsymbol{\beta}_{1j} + \boldsymbol{X}_{2j} \, {}^{\tau}\boldsymbol{\beta}_{2j} \right]$$

where

$\boldsymbol{X}_{1j}$: the vector of covariates for the factor $j =$ (area of residence, education, household wealth, age)

$\boldsymbol{X}_{2j}$: the vector of covariates for the interaction between the factor j and time

${}^{\boldsymbol{i}}\boldsymbol{\beta}_{1j}$: the vector of coefficients for the factor j for the parameter $i = (\mu, \sigma, \nu, \tau)$

${}^{\boldsymbol{i}}\boldsymbol{\beta}_{2j}$: the vector of coefficients for the interaction between the factor j and time for the parameter i

Time was centered at July 1, 2010 so that coefficient estimates for the factor j represent values of the BCPE parameters as of mid-year 2010. The coefficients for the interaction terms represent annual rates of change in the BCPE parameters. After fitting the model separately for country and sex, values of 4 parameters were estimated for each combination of year (2005 or 2010), age group, area of residence, education, and wealth quartile. Then, BMI distribution curves and prevalence of 4 BMI categories were estimated.

Sampling weights and clustering within primary sampling units (PSUs) were incorporated in all analyses. Detailed description is seen in the previous chapter (2.2.6 Analysis methods). For variance estimation, a bootstrapping technique was used. The model did not converge in 100 iterations with some bootstrap samples (41 (2.0%) and 7 (0.3%) out of 2,000 samples from the Mexican male and female data, respectively; 35 (1.8%) and 164 (8.2%) from the Peruvian male and female data, respectively). For these samples, an average of estimates from the last 50 iterations was used as the final estimate after verifying that the estimates of each iteration were not diverging.

3.3 Results

3.3.1 Study population

The distributions of characteristics of the study populations (Table 3-2) were very similar to those observed in the previous objective. This was expected since the differences of the study population

Table 3-2 Demographic characteristics of the study populations (adults aged 20-59 years)*

	Mexico (n = 57,414)				Colombia (n = 135,403)				Peru (n = 30,811)			
	Men (n = 23,252)		Women (n = 34,162)		Men (n = 55,789)		Women (n = 79,614)		Men (n = 13,861)		Women (n = 16,950)	
	2006	2012	2006	2012	2005	2010	2005	2010	2007-8	2012-3	2007-8	2012-3
Number of observations												
Number of records	10,614	12,638	16,358	17,804	21,657	34,132	34,219	45,395	7,083	6,778	8,318	8,632
Weighted counts (000)	26,311	30,648	26,775	30,616	11,137	12,328	11,307	12,537	7,074	7,200	6,943	7,058
Age (%)												
20-29	32	32	30	29	34	32	32	30	32	27	30	25
30-39	28	26	30	30	27	26	27	26	27	25	30	26
40-49	24	24	24	24	23	24	24	25	27	27	26	27
50-59	17	18	16	18	16	17	17	19	14	21	15	22
Type of residence (%)												
Urban	79	79	77	79	74	74	77	78	68	74	71	77
Rural	21	21	23	21	26	26	23	22	32	26	29	23
Educational attainment (%)												
Primary	40	30	49	35	40	36	39	34	25	21	36	32
Secondary	45	51	40	48	42	44	42	43	45	43	37	36
Higher	15	19	11	17	18	20	18	23	30	36	27	32
Household wealth quartile (%)												
Lowest	22	23	24	23	24	25	20	21	21	24	20	21
Lower	25	25	25	25	26	26	26	25	25	26	23	25
Higher	27	26	26	26	25	25	27	27	28	25	29	27
Highest	27	27	25	26	25	24	27	27	26	26	28	27

* All the numbers were calculated using sampling weights except the number of records.

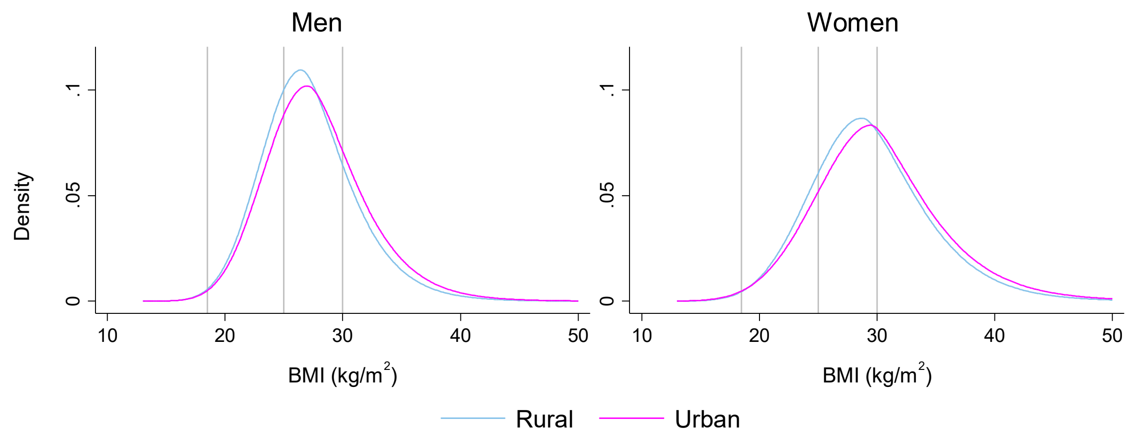
were the exclusions of age group 60-69 years and of those with missing values for education or household wealth that were small in proportion (less than 3.6% among those with valid BMI, Table 3-1). Between two survey years, the proportions of age groups and place of residence (urban or rural) did not seem differ in Mexico and Colombia; it seemed that the population aged, and the urban population increased in Peru. In all the 3 countries, the populations had more educational attainment in the second survey as compared to the first survey.

3.3.2 Estimated BMI distributions in 2010

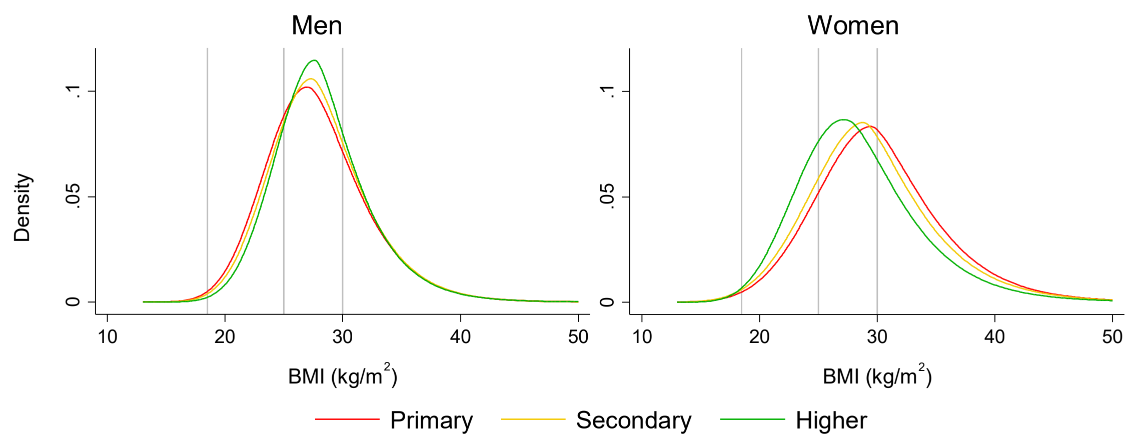
Estimated BMI distributions and prevalence of 4 BMI categories for the year 2010 are presented in Figure 3-1 and 3-2, and their differences between the years 2005 and 20010, in Figure 3-3 and 3-4. Estimated parameters are presented in Appendix F. Estimated BMI distributions were similar

Figure 3-1a Estimated BMI distributions by geographic and socioeconomic factors for year 2010, Mexico (among populations aged 40-49 years)

Place of residence (among population with primary or less education in the lowest wealth quartile)



Education (among urban population in the lowest wealth quartile)



Household wealth (among urban population with primary or less education)

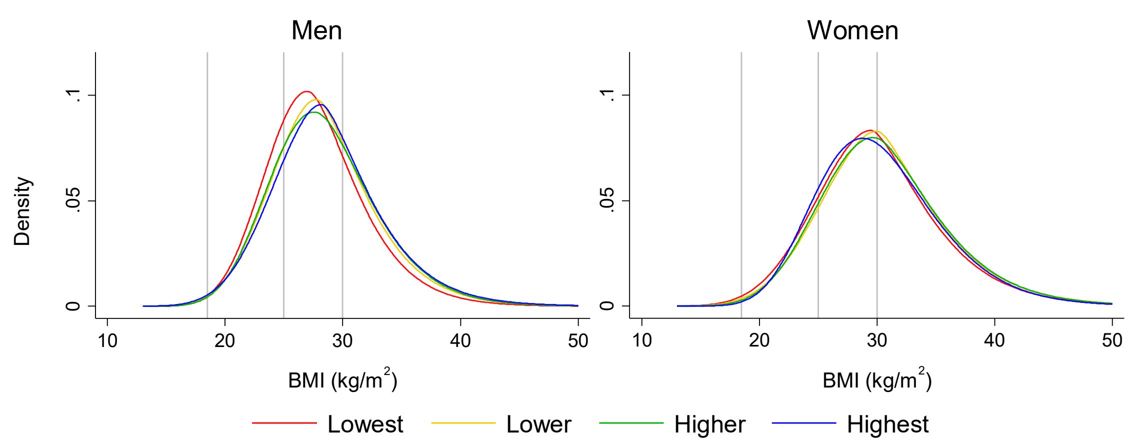
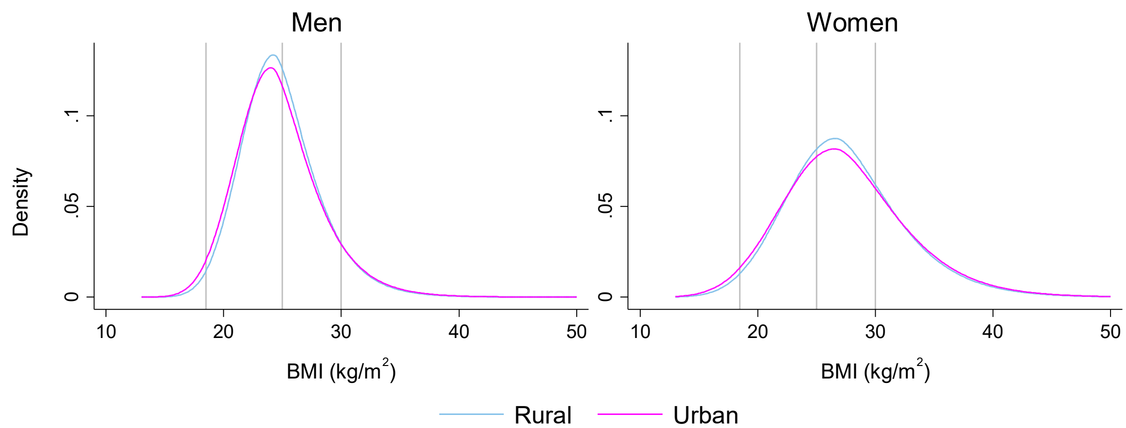
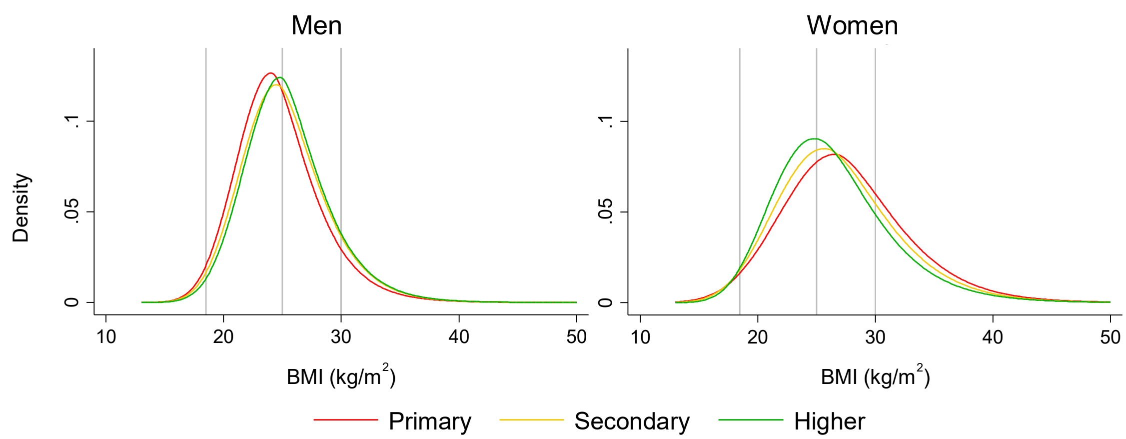


Figure 3-1b Estimated BMI distributions by geographic and socioeconomic factors for year 2010, Colombia (among populations aged 40-49 years)

Place of residence (among population with primary or less education in the lowest wealth quartile)



Education (among urban population in the lowest wealth quartile)



Household wealth (among urban population with primary or less education)

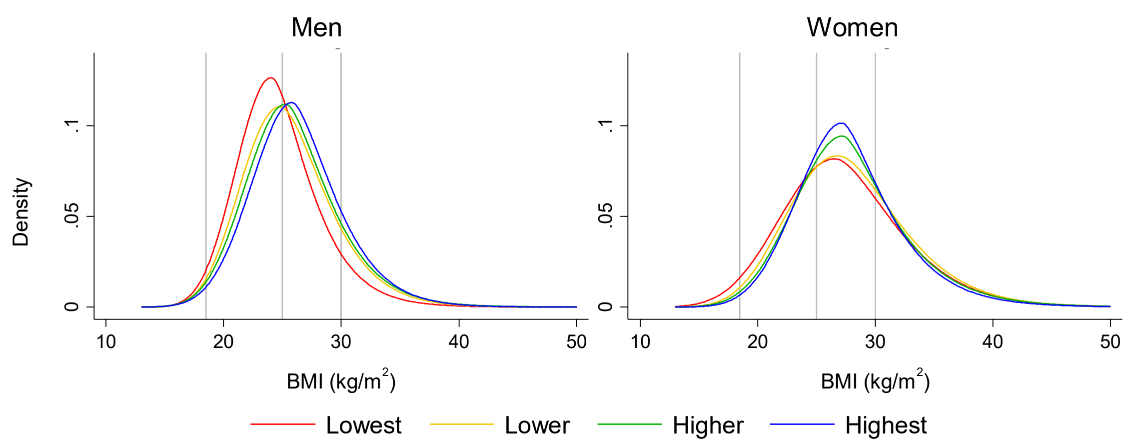
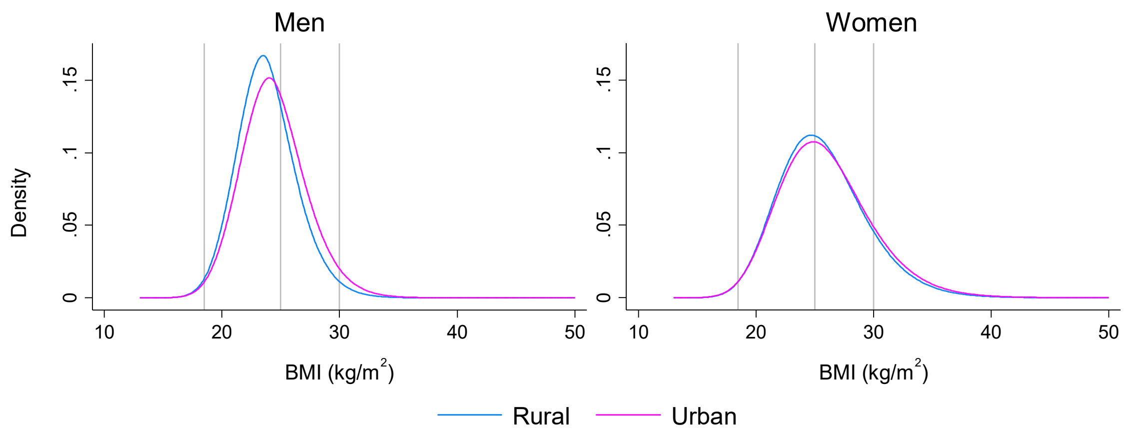
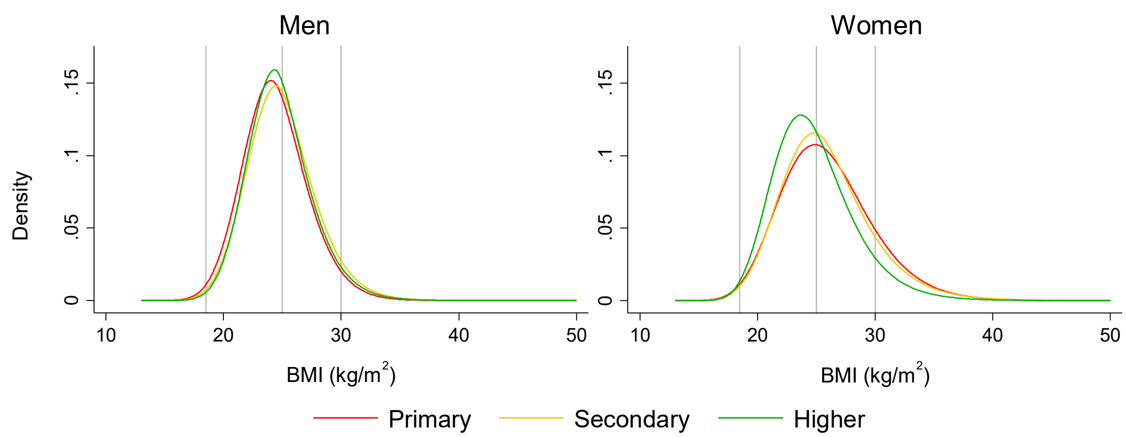


Figure 3-1c Estimated BMI distributions by geographic and socioeconomic factors for year 2010, Peru (among populations aged 40-49 years)

Place of residence (among population with primary or less education in the lowest wealth quartile)



Education (among urban population in the lowest wealth quartile)



Household wealth (among urban population with primary or less education)

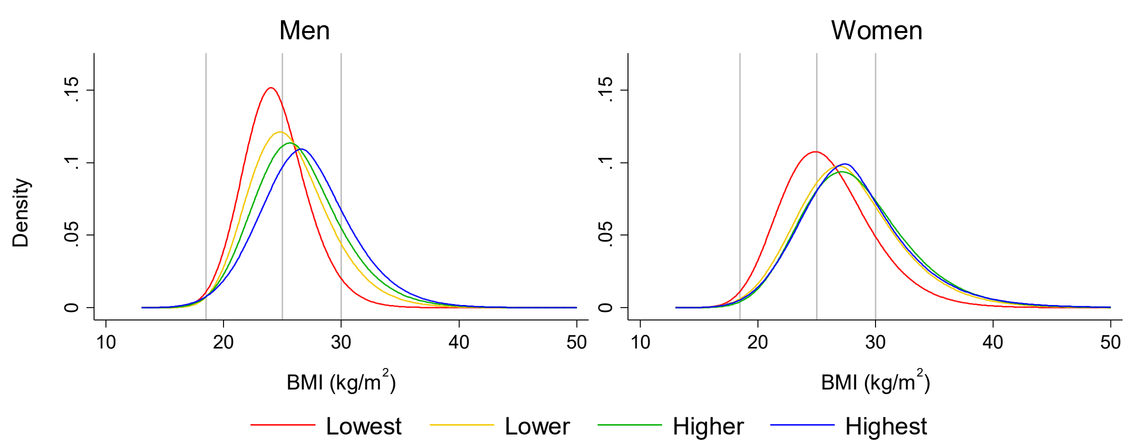
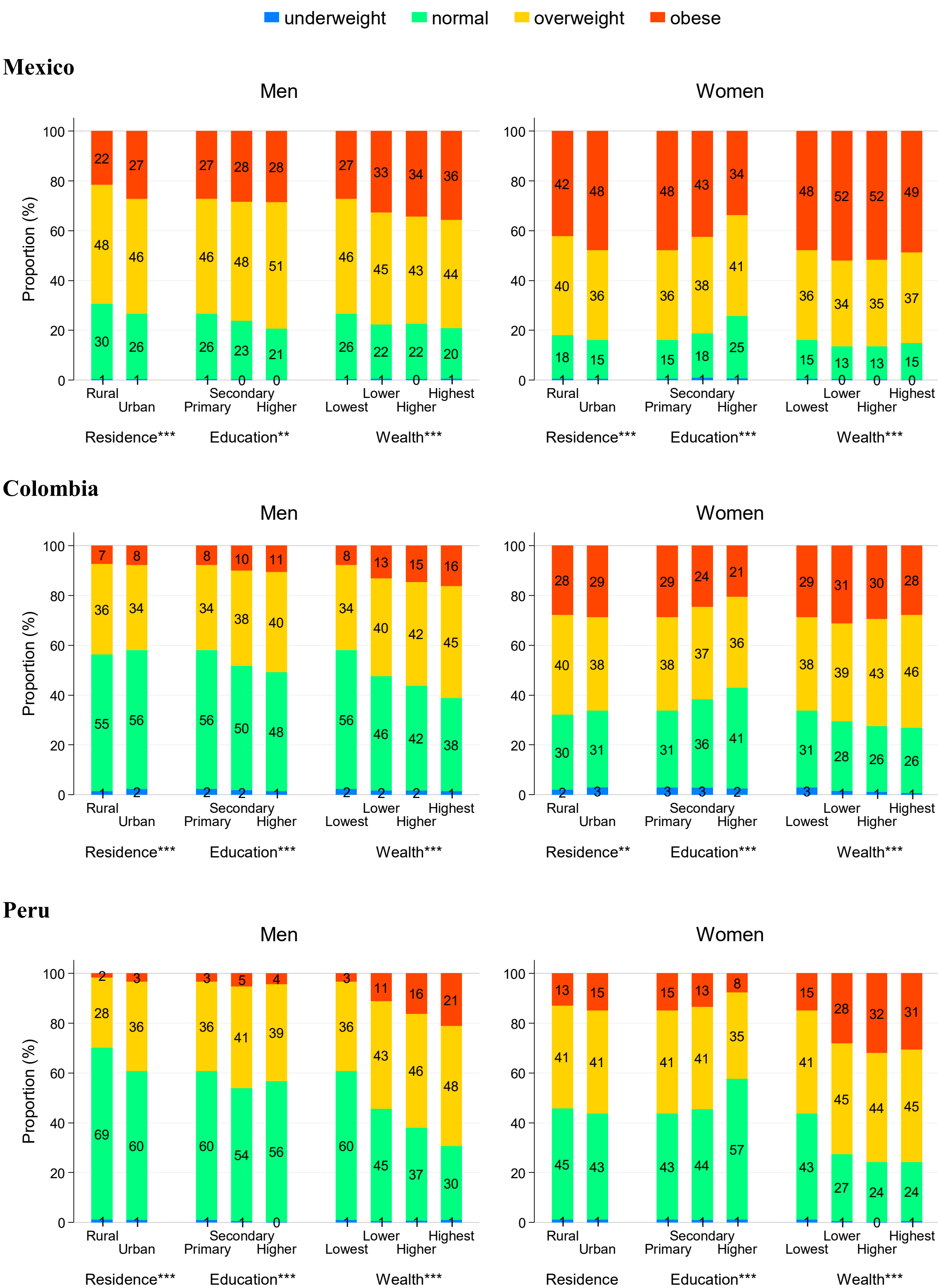


Figure 3-2 Estimated prevalence of 4 BMI categories by geographic and socioeconomic factors for year 2010 (among populations aged 40-49 years)*†



* Regarding the reference groups of each category,

- The prevalence shown under the “residence” present estimated prevalence among people with primary education or less in the lowest wealth quartile;
- The prevalence shown under the “education” present estimated prevalence among urban people in the lowest wealth quartile; and
- The prevalence shown under the “wealth” present estimated prevalence among urban people with primary or less education.

† * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$ in the test for the difference in distributions across categories from $4(k_j - 1)$ *df* likelihood ratio tests. The null hypothesis is: $\beta_{1jk}^{\mu} = 0, \beta_{1jk}^{\sigma} = 0, \beta_{1jk}^{\nu} = 0, \beta_{1jk}^{\tau} = 0, k = 2, 3, \ldots, k_j$ where β_{1jk}^{i} is the coefficient for the category k of the factor $j =$ (place of residence, education, household wealth, age) for the parameter $i = (\mu, \sigma, \nu, \tau)$. See Appendix F for actual p-values.

among adjacent groups but different slightly in their location, scale, and slope because age was a part of the covariates. For this, estimates values for people aged 40-49 years, who were one of the age groups with the highest overweight and obesity prevalence, were selected for the presentation in this chapter.

Education was negatively associated with BMI after adjustment for other covariates in women in all the 3 countries whereas it was somewhat positively associated in men. There was a clear shift of BMI distributions to the left by education in women (Figure 3-1a, -1b, -1c, middle row, right; global test for difference in distributions: $p < 0.001$ in Mexican, Colombian, and Peruvian women). For example, among Mexican women aged 40-49 years in the lowest wealth quartile in urban areas, overweight and obesity prevalence was 84.0% (95% percentile confidence interval: 82.3, 85.6) among those with primary education or less as compared to 74.3% (70.8, 77.7) among those with higher education (Figure 3-2, top right).

Higher wealth was positively associated with BMI in men whereas it was less apparent in women after adjustment for other covariates. There was a clear shift of BMI distribution to the right by wealth especially in Colombian and Peruvian men (Figure 3-1b, -1c, bottom left; global test for difference in distributions: $p < 0.001$ in Mexican, Colombian, and Peruvian men). For example, among Peruvian men aged 40-49 with primary education or less in urban areas, overweight and obesity prevalence was 39.2% (35.0, 43.2) among those in the lowest wealth quartile as compared to 69.2% (66.0, 72.5) among those in the highest wealth quartile (Figure 3-2, bottom left). In women, positive associations were seen among the lowest and lower wealth quartile groups, but the associations became weak (in Colombia and Peru) or reversed (in Mexico) among the higher wealth groups (Figure 3-1a, -1b, -1c, bottom right; Figure 3-2, right).

Urban residence was associated with higher BMI among Mexican men and women, and Peruvian men (Figure 3-1a, -1c, top row; Figure 3-2; global test for difference in distributions: $p < 0.001$ for all the 3 groups). For example, among Peruvian men aged 40-49 with primary education or less in the lowest wealth quartile, overweight and obesity prevalence was 39.2 % (35.0, 43.2) among those in urban areas as compared to those in in rural area 30.0% (27.1, 32.9). In Colombia, urban residence was negatively associated; however, the difference did not seem clinically important (Figure 3-1b, top row; Figure 3-2; global test for difference in distributions: $p < 0.001$ for men and $p < 0.01$ for women).

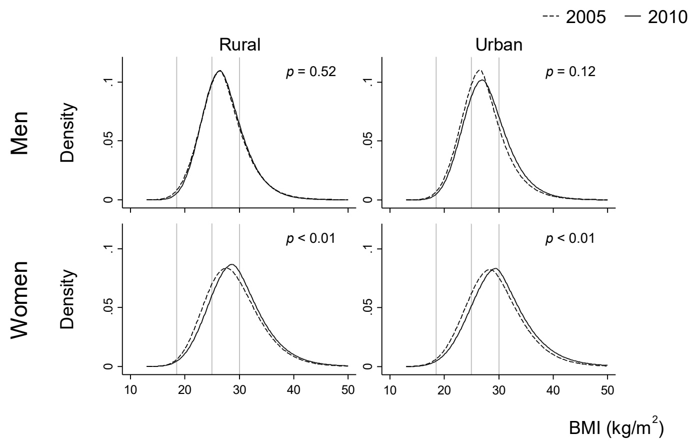
3.3.3 Estimated differences in BMI distributions between 2005 and 2010

Regarding the rates of change in BMI distributions, lower wealth was associated with higher rates of change in women's BMI distributions in Mexico and Colombia between 2005 and 2010 (Figure 3-3a, -3b, bottom row; global tests for difference in distributions across wealth quartiles: $p = 0.03$, $p < 0.001$, respectively). A similar association was seen in Peruvian women but was not significant ($p = 0.10$). For example, among Colombian women aged 40-49 years with primary education in urban areas, overweight and obesity prevalence increased by 7.9 percentage point (pp) (5.3, 10.4) among those in the lowest wealth quartile as compared to 3.6 pp (1.4, 5.9) among those in the highest wealth quartile (Figure 3-4, bottom right). A gradient was seen in association between wealth and rates of change in overweight and obesity prevalence across the 3 countries (Figure 3-4).

Apart from this, no clear patterns were observed across the countries. Among Colombian men, those with secondary education or higher tended to increase overweight and obesity prevalence faster than those with primary education or less (Figure 3-4, middle left).

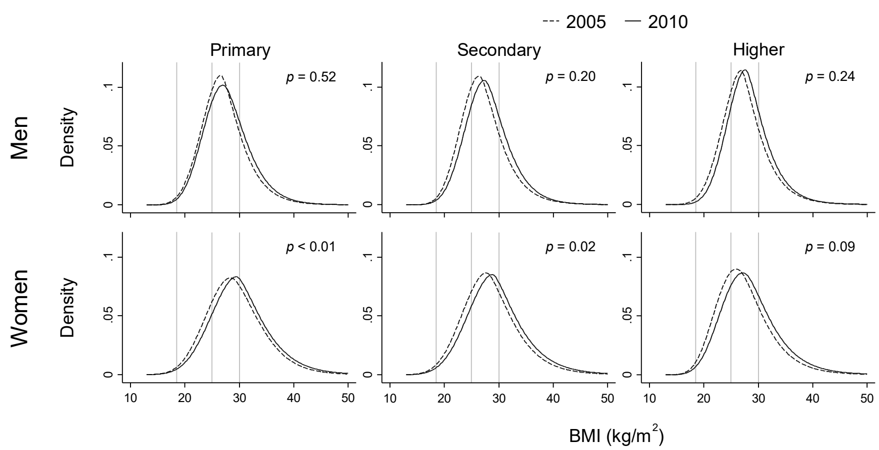
Figure 3-3a Estimated BMI distributions by geographic and socioeconomic factors for years 2005 and 2010, Mexico (among populations aged 40-49 years)

Place of residence (among population with primary or less education in the lowest wealth quartile)



The p-values is from 4 df likelihood ratio tests. The null hypothesis is: $\beta^{\mu}_{2jk} + \beta^{\mu}_{24k} = 0$, $\beta^{\sigma}_{2jk} + \beta^{\sigma}_{24k} = 0$, $\beta^{\nu}_{2jk} + \beta^{\nu}_{24k} = 0$, $\beta^{\tau}_{2jk} + \beta^{\tau}_{24k} = 0$, $k = 1, 2, \ldots, k_j$, where β^{i}_{2jk} is the coefficient for the interaction term between the category k of the factor $j =$ (place of residence, education, household wealth, age) and time for the parameter $i = (\mu, \sigma, \nu, \tau)$.

Education (among urban population in the lowest wealth quartile)



Household wealth (among urban population with primary or less education)

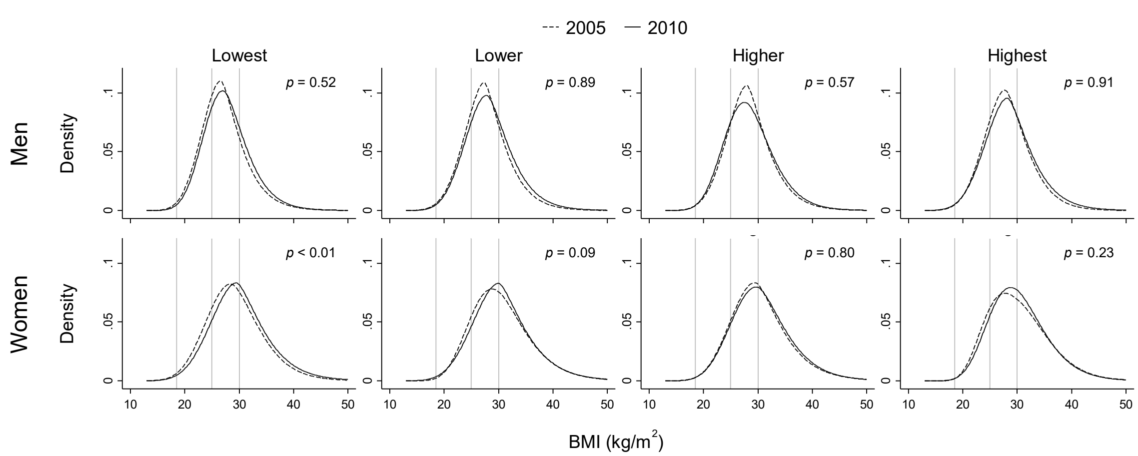
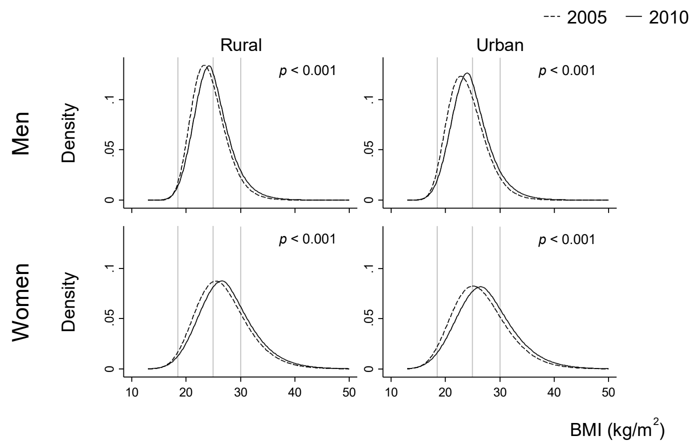


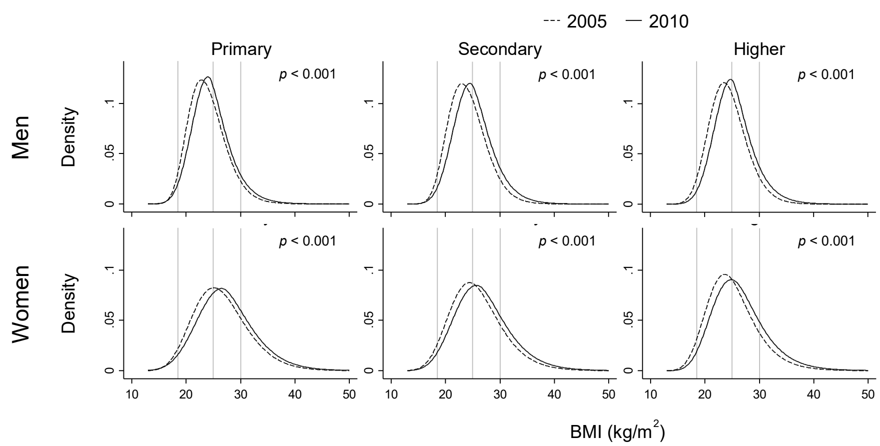
Figure 3-3b Estimated BMI distributions by geographic and socioeconomic factors for years 2005 and 2010, Colombia (among populations aged 40-49 years)

Place of residence (among population with primary or less education in the lowest wealth quartile)



The p-values is from 4 df likelihood ratio tests. The null hypothesis is: $\beta^{\mu}_{2jk} + \beta^{\mu}_{24k} = 0$, $\beta^{\sigma}_{2jk} + \beta^{\sigma}_{24k} = 0$, $\beta^{\nu}_{2jk} + \beta^{\nu}_{24k} = 0$, $\beta^{\tau}_{2jk} + \beta^{\tau}_{24k} = 0$, $k = 1, 2, \ldots, k_j$, where β^{i}_{2jk} is the coefficient for the interaction term between the category k of the factor $j =$ (place of residence, education, household wealth, age) and time for the parameter $i = (\mu, \sigma, \nu, \tau)$.

Education (among urban population in the lowest wealth quartile)



Household wealth (among urban population with primary or less education)

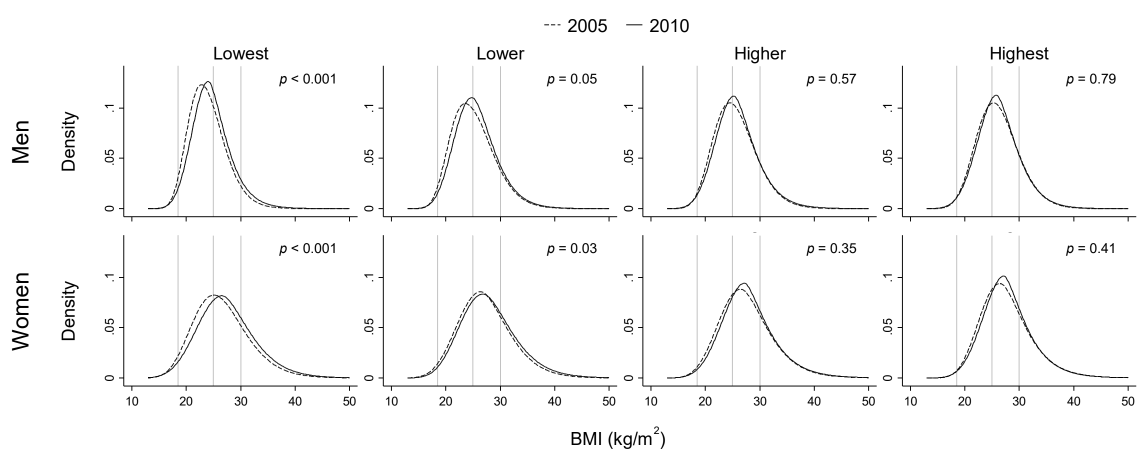
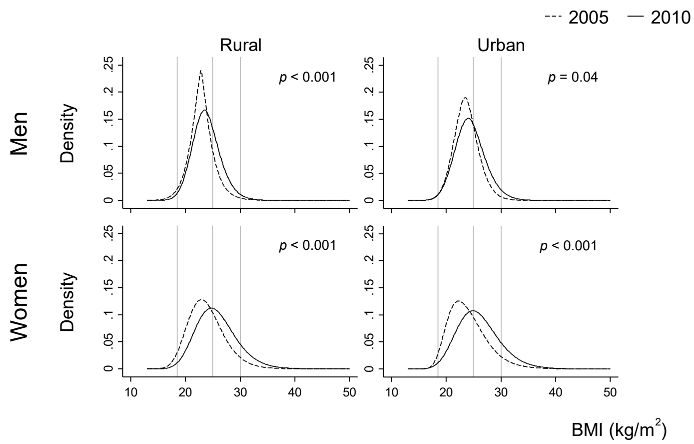


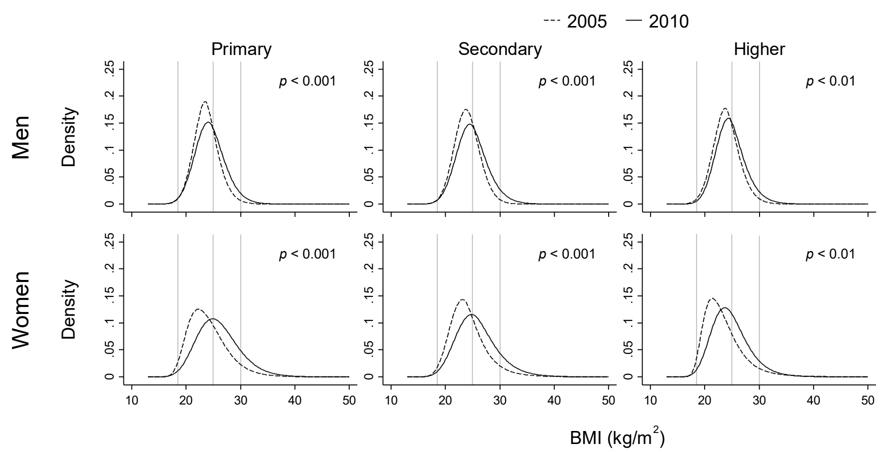
Figure 3-3c Estimated BMI distributions by geographic and socioeconomic factors for years 2005 and 2010, Peru (among populations aged 40-49 years)

Place of residence (among population with primary or less education in the lowest wealth quartile)



The p-values is from 4 df likelihood ratio tests. The null hypothesis is: $\beta^{\mu}_{2jk} + \beta^{\mu}_{24k} = 0$, $\beta^{\sigma}_{2jk} + \beta^{\sigma}_{24k} = 0$, $\beta^{\nu}_{2jk} + \beta^{\nu}_{24k} = 0$, $\beta^{\tau}_{2jk} + \beta^{\tau}_{24k} = 0$, $k = 1, 2, \dots, k_j$, where β^{i}_{2jk} is the coefficient for the interaction term between the category k of the factor $j =$ (place of residence, education, household wealth, age) and time for the parameter $i = (\mu, \sigma, \nu, \tau)$.

Education (among urban population in the lowest wealth quartile)



Household wealth (among urban population with primary or less education)

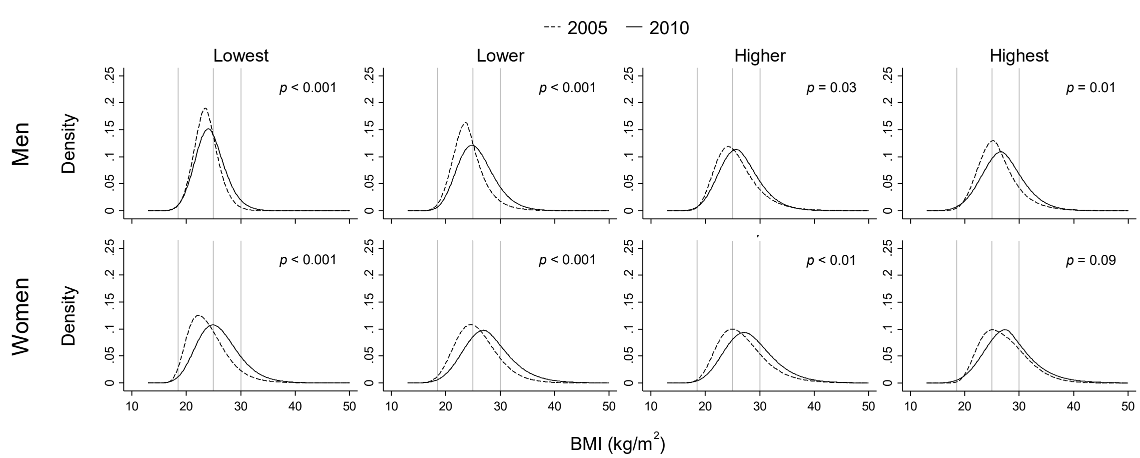
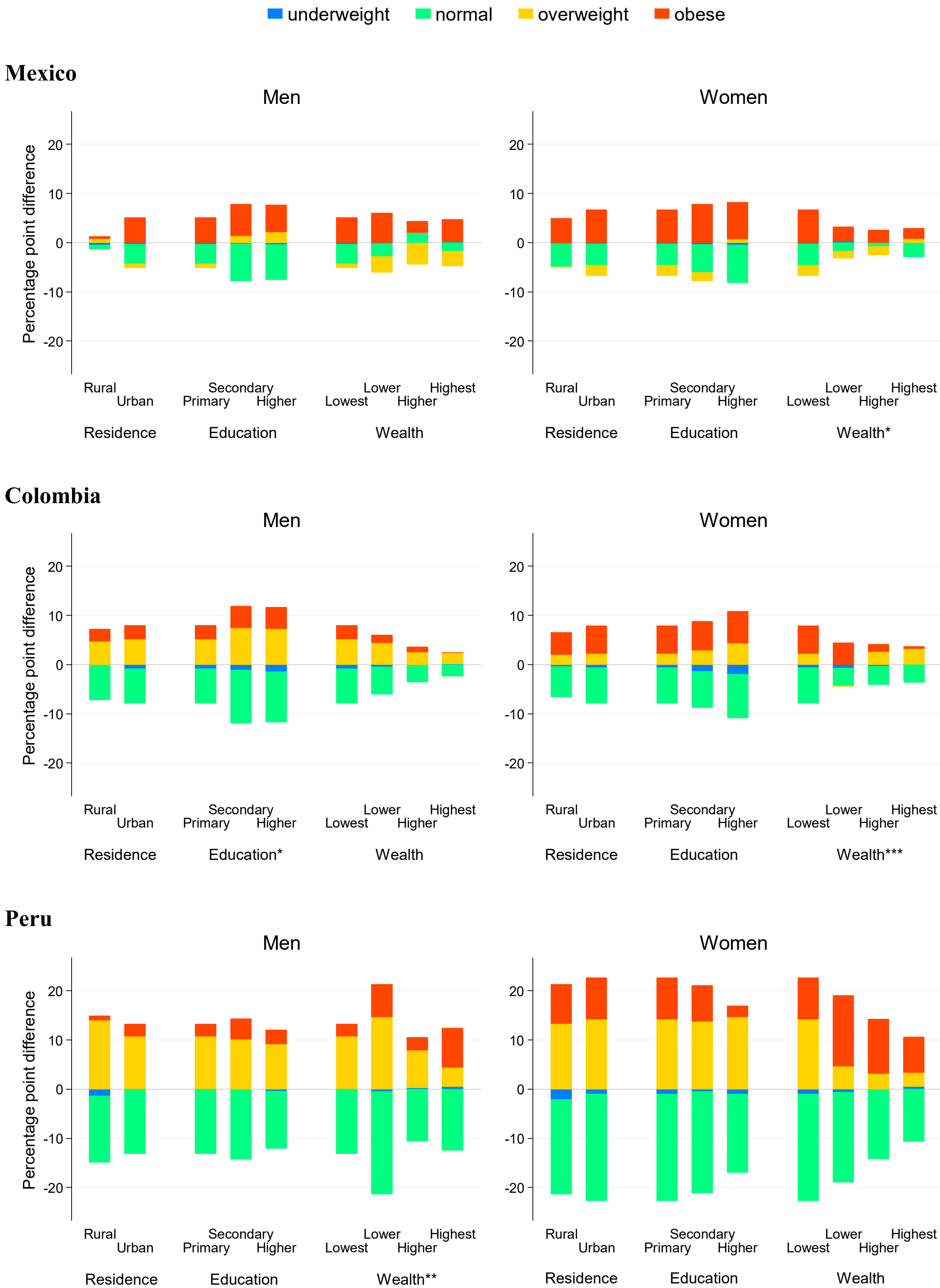


Figure 3-4 Estimated change in prevalence of 4 BMI categories by geographic and socioeconomic factors between years 2005 and 2010 (among populations aged 40-49 years)*†



* Regarding the reference groups of each category,
  - The prevalence shown under the "residence" present estimated prevalence among people with primary education or less in the lowest wealth quartile;
  - The prevalence shown under the "education" present estimated prevalence among urban people in the lowest wealth quartile; and
  - The prevalence shown under the "wealth" present estimated prevalence among urban people with primary or less education.

† * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$ in the test for the difference in change in distributions between 2005 and 2010 across categories from $4(k_j - 1)$ *df* likelihood ratio tests. The null hypothesis is: $\beta^{\mu}_{2jk} = 0, \beta^{\sigma}_{2jk} = 0, \beta^{\nu}_{2jk} = 0$, $\beta^{\tau}_{2jk} = 0, k = 2, 3, \dots, k_j$ where β^{i}_{2jk} is the coefficient for the interaction term between the category k of the factor $j =$ (place of residence, education, household wealth, age) and time for the parameter $i = (\mu, \sigma, \nu, \tau)$. See Appendix F for actual p-values.

3.4 Discussion

3.4.1 Main findings

Common results across the 3 countries were: [1] education was negatively associated with BMI in women whereas it was somewhat positively associated in men; [2] wealth was positively associated with BMI in men whereas it was associated positively only among those in the lowest and lower wealth quartiles in women; and [3] rates of change in overweight and obesity were lower among women with more wealth (or rates of change was higher among those with less wealth), after adjustment for other covariates. Hence, there was a tendency that women were more sensitive to weight or BMI as their socioeconomic status increased whereas this relationship was not present for men.

These observations had similarities with the findings of previous studies. Our studies, which included 3 upper-middle-income countries, were consistent with the observations from middle- and high-income countries, where socioeconomic status was negatively associated with BMI in women, and the associations were mixed and inconsistent in men (McLaren 2007, Pampel et al. 2012, Dinsa et al. 2012). A study from Brazil, another upper-middle-income country in the region, reported similar sex-differences in associations between socioeconomic status and BMI (Monteiro et al. 2007). Their analyses of 3 national surveys revealed that positive associations between family income and obesity prevalence were observed across 3 surveys between 1975 and 2003 in men. On the other hand, an inverted U-shaped association was seen with its peak at the fourth quintile in women in the first survey.

Obesity prevalence in women increased more rapidly in lower income groups afterwards, and the association became flat or somewhat negative in the last survey.

More specifically, our findings about negative associations between education and BMI, and less apparent positive association between wealth and BMI among women than men were consistent with the results of a multi-county study based on the DHS data from 36 low- and middle-income countries collected in 1991-2008 (Mamun and Finlay 2014). The authors reported that positive associations of wealth and education with BMI were weakening over time. Their findings implied that weakened positive associations of socioeconomic factors with BMI are expected as a country develops, which were actually observed among women in Mexico, Colombia, and Peru.

Negative associations between household wealth and rates of change in overweight and obesity prevalence among women were consistent with the findings of a multi-country study based on mainly the DHS data of 2 time points from 39 low- and middle-income countries collected in 1991-2008 (Jones-Smith et al. 2012). The authors reported that the higher the GDP per capita, the more rapidly the lowest wealth group increased in overweight and obesity prevalence than the highest wealth group.

The observed sex differences may be due to differences in perception in obesity between men and women, and cultural norms that exist. As a national economy progresses, people with higher socioeconomic status may start having greater demand for health and thinness whereas people with lower socioeconomic status may continue having demand for calories (Goryakin et al. 2014). However, women may be more sensitive to and concerned about weight as compared to men due to existing cultural norms in which slim women are recognized as a symbol of success (Jacoby et al. 2003).

Difference in BMI distributions were observed between urban and rural areas in Mexico, and in Peru among men in 2010, after adjustment for individual socioeconomic status measured by education and household wealth. This implies that there was still a difference in energy balance between urban and rural area that are not explained by individual education and household wealth. It might be natural to think that the level of energy consumption is greater and the level of physical activity is less in urban areas than in rural areas on average. An interesting observation was that such an urban-rural difference was not observed in Colombia. In case of women, there was little difference in overweight and obesity prevalence between urban and rural areas before adjustment (data not shown).

Whereas there is correlation between education and wealth in general, we observed somewhat different associations of education and wealth with BMI distributions as discussed above. Inconsistent associations of education and wealth with overweight and obesity have been reported in several studies in Peru (Jacoby et al. 2003, Poterico et al. 2012, Álvarez-Dongo et al. 2012, Quispe et al. 2016), in Brazil (Monteiro et al. 2001), and in a multi-country study (Pampel et al. 2012, Jones-Smith et al. 2012). Education is a knowledge-related asset of a person that is usually completed in young adulthood or before, and household assets and income are available resources at the time of survey (Galobardes et al. 2006, Quispe et al. 2016). Hence, education and wealth reflect different aspects of socioeconomic status and may have different effects on overweight and obesity to a certain extent.

3.4.2 Consistencies and discrepancies with country-specific studies

A similar study was conducted using the same data of Mexico (Quezada and Lozada-Tequeanes 2015). The authors reported: a positive association of both education and household wealth with overweight and obesity in men; and a negative association of education and an inverted U-shaped association of wealth in women. These findings were consistent with ours. However, they reported higher increases in obesity prevalence among women in middle- to upper-middle wealth, which was different from our finding that obesity prevalence increased more among women in lower wealth groups. The main causes of the discrepancy would be differences in methods to construct wealth indices and to calculate prevalence. Quezada and Lozada-Tequeanes constructed absolute wealth indices, which are known as comparative wealth indices because they are comparable across surveys (Rutstein and Staveteig 2014). On the other hand, we constructed relative wealth indices that are relative scores at each particular point in time. For example, two households with the same dwelling characteristics and household assets in two surveys have the same absolute wealth index score but different relative wealth index scores. Another main cause could be that they used predictive margins to calculate the prevalence whereas our study used covariate-specific prevalence. Other possible causes include the difference in variables used to construct wealth index and the difference in included covariates in the regression models. Another study that was based on 4 consecutive survey data of women of reproductive age conducted between 1995 and 2010 reported an inverse association of education with obesity prevalence in urban areas but not in rural areas (Perez Ferrer et al. 2014). Since we assumed equal effects of socioeconomic factors in urban and rural areas, further investigation is necessary about this point.

A similar study was also conducted using the same data of Colombia. The authors reported a positive association between wealth and obesity and the fastest increases in obesity prevalence among people of the lowest wealth and among those living in urban areas (Kasper et al. 2013). While the results about wealth were consistent, those for urbanicity were different. Our study indicated that there was little difference in rates of change between urban and rural areas. The discrepancy might be due to the differences in covariates included in the models and the fact that sex was included as a covariate without its interaction with other variables in their models. In our study, regression analyses were conducted separately by sex.

Our findings were consistent with previous studies in Peru. Health examination surveys conducted in 6 cities in Peru between 1998 and 2000 found that years of education was associated with overweight and obesity positively in men but negatively in women (Jacoby et al. 2003). A cross-sectional study of the National Health Survey (ENAHO) conducted in 2009-2010 reported higher prevalence of overweight and obesity in women, urban areas, and people not poor (Álvarez-Dongo et al. 2012). Another cross-sectional study using DHS conducted in 2008 reported a positive association between wealth and obesity (Poterico et al. 2012). However, the authors reported a stronger association in rural areas than urban areas, which implied an interaction effect of place of residence. This requires further investigation as discussed earlier. A cross-sectional study using the baseline data of CRONICAS cohort study reported that family income was associated positively with obesity whereas education was insignificantly positively associated although they pooled men and women in analyses (Quispe et al. 2016).

3.4.3 Implications for public health practice

Differences in magnitude and rates of change in BMI distributions were observed by geographic and socioeconomic factors, and by sex. As the country's economy has progressed, acknowledging that there are substantial differences in obesity situation and surrounding conditions among the assessed 3 countries, it seemed that women of high socioeconomic status have already started to enter the next stage of the nutritional transition, from the stage where obesity emerges and increases to the next stage where reduction of body fat occurs (Popkin and Gordon-Larsen 2004). On the other hand, this transition did not seem occurring yet among men.

These imply the necessity of tailoring relevant health policies and programs to reflect observed patterns of BMI distributions if and where appropriate. For example, health educational program could be modified in terms of how to reach the target populations (e.g., TV programs, screening at health facilities, and workplace programs, with effective language) and what messages will be given according to the problems they face (i.e., already affected or being affected). Population level interventions, which could work better for target populations than other populations, might be sought. For instance, there were some reports that taxation on sugar sweeten beverages would reduce the consumptions of such beverages especially among poor (Claro et al. 2011, Colchero et al. 2015). Hence, after capturing trends of obesity distributions and investigating further about determinants of increases and decreases in obesity, a review of current health policies and programs would be necessary (other aspects of implications were discussed in the previous chapter, 2.4.3 Implication for public health practice).

3.4.4 Strengths and limitations

There are some strengths of our study, which were fully discussed in the previous chapter (2.4.4 Strengths and weaknesses). Among those, it should be highlighted that our study was based on survey data that included both men and women. As seen in the results, some associations differed by sex. Previous major studies on geographic and socioeconomic factors in low- and middle-income countries were based solely on data from women of reproductive age (Popkin et al. 2012), and our study provides additional information by including male populations in the analyses. Also, our study assessed both magnitude and rates of change in BMI distributions to identify subpopulations that currently suffer from the problem of the overweight and obesity as well as those being more affected as compared to other groups.

There are some limitations of our study as well. Some previous studies have suggested interactions between place of residence (urban-rural area), education, and wealth (Subramanian et al. 2011, Poterico et al. 2012, Perez Ferrer et al. 2012, Aitsi-Selmi et al. 2014, Ouyang et al. 2015). However, the interactions were not considered in our study since our focus was to assess the averaged effects of each factor on the populations. Hence, stratified analysis or regression analysis including interaction terms would be the next analysis to be conducted. Computational time in estimating variances with bootstrap samples could be an issue when it comes to apply the method used in our study for other data. It took about one to two days to obtain bootstrapped samples for each. However, on-going improvement in technology would solve this.

Chapter 4 Obesity prediction by modeling BMI distribution using national health survey data

4.1 Objective

The objective of this component of the study is to assess obesity prediction performance by modeling the entire BMI distributions.

4.2 Methods

4.2.1 Data sources

Data at 4-5 time points from nationally representative health and nutritional household surveys conducted between 1998 and 2014 in Mexico, Colombia, and Peru were used. Since data at 4 or more time points were available only for women of reproductive age, this study was limited to them.

For Mexico, data of the national surveys conducted in 1988 (National Nutritional Survey or Encuesta Nacional de Nutrición, ENN), in 2000 (National Health Survey or Encuesta Nacional de Salud, ENS), and in 2006 and 2016 (National Health and Nutritional Survey or Encuesta Nacional de Salud y Nutrición, ENSANUT) were used (Sepúlveda-Amor et al. 1990, Valdespino et al. 2003, Olaiz et al. 2003, Olaiz-Fernández et al. 2006, Shamah-Levy et al. 2007, Gutiérrez et al. 2013). The last two surveys were the same data sources used in the previous two chapters. ENN 1988 measured children under 5 years of age and women of reproductive age; and the rest of 3 surveys measured all the selected household members regardless of sex and age. The National Nutritional Survey conducted in 1999 was not included due to unavailability of its complete data. The data were obtained from the website of the Instituto Nacional de Salud Publica (INSP), Mexico, after authorization (INSP 2009a, 2009b, 2012a, 2012b).

For Colombia, data from the Demographic and Health Survey (Encuesta Nacional de Demografía y Salud, ENDS) conducted in 1995, 2000, 2005 and 2010 were used (Profamilia and Macro International 1995a, 2000a, 2005a, Profamilia and IFC Macro 2011a). The last two surveys were the same data sources used in the previous two objectives. ENDS 1995 and 2000 measured children under 5 years of age and women of reproductive age. Among the women, anthropometry was measured for those who had a birth history within 5 years from the survey. The recent 2 surveys collected anthropometric measurements from all household members aged 64 years or below for both sexes, regardless of birth history for women. The data were obtained from the website of the DHS Program after authorization (Profamilia and Macro International 1995b, 2000b, 2005b, Profamilia and IFC Macro 2011b).

For Peru, data from the Demographic and Health Survey (Encuesta Demographica y de Salud Famliar, ENDES) conducted in 2000, 2005, and every year from 2006 to 2014 were used (INEI and Macro International 2001a; INEI and ORC Macro 2007, 2009, 2010a, 2011a; INEI 2012e, 2013c, 2014a, 2015a). Data of ENDES, not those of ENAHO that were used in the previous chapters, were used for this component of the study since ENAHO started including anthropometric measurement recently and did not provide enough time points of data. Peru started implementing the DHS as a continuous survey since 2003, and 3 rounds of annual surveys have been conducted so far (2003-2008, 2009-2011, and 2012-2014) (INEI and ORC Macro 2007, 2009, 2012e, 2015a, Rutstein and Way 2014). In the continuous survey, primary sampling units (PSUs) are selected and distributed across the years in a given round so that data could produce selected national indicators annually with a certain precision. ENDES surveyed children under 5 years of age and women of reproductive age. For women, anthropometry was measured regardless of their birth history. In the first round (2003-2008), anthropometric data were collected only in 2005, 2007 and 2008. Data from 2005 and 2007 surveys were pooled in this analysis following the recommendation that data of 2 years should be pooled for anthropometric indicators for their reliable estimates in use of surveys conducted in 2003-2007 (Rutstein and Way 2014). The data were obtained from the website of the DHS Program after authorization (INEI and Macro International 2001b; INEI and ORC Macro 2010b, 2010c, 2011b; INEI 2012f, 2013d, 2014b, 2015b).

All data were from household surveys implemented with stratified multistage cluster sampling. Further details about sampling designs of each survey were described in Appendix C. Similar to the surveys that were used in the previous chapters, primary sampling units (PSUs) were sampled from each stratum within each sub-national unit with probability proportional to size of dwellings or populations in all 3 countries. Stratification units, from which PSUs were sampled, consisted of combinations of the sub-national unit and the urban-rural area or finer area category. The ultimate sampling unit was dwelling, from which all households were sampled. All household members or some of the household members were followed with household and individual questionnaires, and anthropometric measurement. Weight and height were measured using preset standardized procedures by trained survey teams.

Household response rates ranged from 87 to 99% (Table 4-1). Among women of reproductive age (age 15-49 years) overall response rates ranged from 92 to 98% in Colombia and Peru where the data were available. About 95% or more interviewed women were measured for anthropometry in almost all surveys.

4.2.2 Inclusion and exclusion criteria

The study populations for this objective were non-pregnant women aged 20-49 years. Although all the 3 countries started collecting anthropometric data including men and postmenopausal women, it was recent, and only 2 time points of data were available for these subpopulations. For this, the survey data collected from women of reproductive age, which were available at 4 or more time points for more than decades in the countries, were used. In case of Colombia, the study population was further limited to non-pregnant women aged 20-39 years who had a birth history within 5 years from the survey. Their earlier 2 surveys in 1995 and 2000 measured anthropometry of women who met this birth history criterion, and the data were limited to this subgroup. The number of such women aged 40 years or above was small, and the sample was also limited to women aged 20-39 years for Colombia. Anthropometric exclusion criteria were the same as the previous objectives. Those with weight < 20 kg, height < 110 cm or > 210 cm, or BMI < 12 kg/m^2 were excluded considered as implausible values. Those with either weight or height measurements only were excluded since BMI could not be calculated. However, those who were excluded from the study due to incomplete or implausible anthropometric values were usually less than 0.5% (Table 4-1).

Table 4-1 Response rates and sample sizes

Mexico

Results	Survey year			
	1988	2000	2006	2012
Household response rate[*†]	--	--	--	87%
Women response rate	--	--	--	--
Interviewed women[§] (age 20-49)	12,189	22,293	18,040	15,713
Women used in analyses[§]	10,744	20,079	14,044	14,601
Anthropometry response rate[§‖]	95%	98%	81%	96%
Measurement/recording error rate[§¶]	0.5%	1.4%	0.6%	0.1%

Colombia

Results	Survey year			
	1995	2000	2005	2010
Household response rate[*†]	92%	93%	88%	92%
Women response rate[*‡] (age 15-49)	92%	92%	92%	94%
Interviewed women[§] (age 20-39)	3,271	3,042	9,741	11,824
Women used in analyses[§]	2,828	2,722	8,720	10,376
Anthropometry response rate[§‖]	95%	97%	95%	93%
Measurement/recording error rate[§¶]	0.5%	0.0%	0.1%	0.0%

Peru

Results	Survey year				
	2000	2004-6[#]	2007-8[#]	2009	2014
Household response rate[*†]	98%	99%	99%	99%	98%
Women response rate[*‡] (age 15-49)	95%	98%	98%	98%	97%
	Survey year				
	2000	2005-7	2008	2009	2014
Interviewed women[§] (age 20-49)	22,095	10,142	13,137	19,621	20,521
Women used in analyses[§]	20,168	9,176	11,767	18,025	19,585
Anthropometry response rate[§‖]	96%	95%	95%	97%	99%
Measurement/recording error rate[§¶]	0.2%	0.3%	0.3%	0.3%	0.0%

* Data obtained from the final survey reports or publications.

† $\text{Household response rate} = \dfrac{\text{Number of interviewed households}}{\text{Number of occupied households}}$

‡ $\text{Women response rate} = \dfrac{\text{Number of interviewed women}}{\text{Number of eligible women in the interviewed households}}$

§ Data obtained from the datasets.

‖ $\text{Anthropometry response rate} = \dfrac{\text{Number of women with weight and/or height measurement}}{\text{Number of interviewed women}}$

¶ Measurement/recording error rate

$$= \frac{\text{Number of women with incomplete or implausible weight, height or BMI}}{\text{Number of women with weight and/or height measurement}}$$

\# The data used for the study was of 2005-07 and 2008. However, the household and women response rates come from the published reports, which could not be disaggregated for a specific year. Therefore, the rates of closest report year are presented.

4.2.3 Data preparation

Sampling weights were calibrated before the analyses in the same way as described in the previous chapter (2.2.5 Data preparation). Calibration was done so that: the same proportions of age-specific sums of sampling weights could be maintained before and after the exclusion of records; and the total sums of sampling weights (i.e., total population size) for each survey would match with the standard populations estimated over time with the same estimation method.

4.2.4 Analysis method

For each country, all the survey data before the last survey were used as a training set to construct prediction models. In case of Peru, the data from surveys conducted between 2009 and 2013 were reference points without participating in the construction of the training set nor as the validation set. Three obesity indicators were predicted for each 5-year age group for the time when the last survey was conducted. The predicted obesity indicators were median BMI, overweight and obesity prevalence (BMI $\geq$ 25 kg/m^2), and obesity prevalence (BMI $\geq$ 30 kg/m^2). The sample sizes that were used to construct prediction models were 44,867 for Mexico, 14,270 for Colombia, and 59,136 for Peru.

There were two prediction methods that were assessed. With one method, the entire BMI distributions were modeled, assuming that BMI follows a Box-Cox Power Exponential (BCPE) distribution and using the generalized additive model for location, scale and shape (GAMLSS) (referred as “BCPE method”). Each of the 4 BCPE parameters were fitted as a function of age, time, their higher order terms, and their interaction terms. In a second method, each of the obesity indicators was directly modeled as a function of the same covariates using quantile regression (Koenker and Bassett 1978) to model median BMI and logistic regression to model overweight and obesity prevalence (referred as “direct method”). In both methods, age was included as nominal or ordered 5-year age groups, which made a total of 4 models for comparison. The model equations were as follows:

<u>BCPE-GAMLSS method with age group as a nominal variable</u>

$$BMI \sim BCPE(\mu, \sigma, \nu, \tau)$$

$$\mu = \sum\nolimits_k [\,{}^{\mu}\beta_{1k} age_k + {}^{\mu}\beta_{2k}(age_k \times time) + {}^{\mu}\beta_{3k}(age_k \times time^2)]$$

$$\log \sigma = \sum\nolimits_k [\,{}^{\sigma}\beta_{1k} age_k + {}^{\sigma}\beta_{2k}(age_k \times time) + {}^{\sigma}\beta_{3k}(age_k \times time^2)]$$

$$\nu = \sum\nolimits_k [\,{}^{\nu}\beta_{1k} age_k + {}^{\nu}\beta_{2k}(age_k \times time) + {}^{\nu}\beta_{3k}(age_k \times time^2)]$$

$$\log \tau = \sum\nolimits_k [\,{}^{\tau}\beta_{1k} age_k + {}^{\tau}\beta_{2k}(age_k \times time) + {}^{\tau}\beta_{3k}(age_k \times time^2)]$$

where

age_k: the indicator variable for the 5-year age group k

$time$: the continuous variable for time in year centered at the time of the last survey

${}^{i}\beta_{jk}$: the coefficient for the age category k of the covariate j for the parameter $i = (\mu, \sigma, \nu, \tau)$

BCPE method with age group as an ordinal variable

$$BMI \sim BCPE(\mu, \sigma, \nu, \tau)$$

$$\mu = {}^{\mu}\beta_1 age + {}^{\mu}\beta_2 age^2 + {}^{\mu}\beta_3 time + {}^{\mu}\beta_4 time^2$$
$$+ {}^{\mu}\beta_5(age \times time) + {}^{\mu}\beta_6(age^2 \times time) + {}^{\mu}\beta_7(age \times time^2)$$

$$\log \sigma = {}^{\sigma}\beta_1 age + {}^{\sigma}\beta_2 age^2 + {}^{\sigma}\beta_3 time + {}^{\sigma}\beta_4 time^2$$
$$+ {}^{\sigma}\beta_5(age \times time) + {}^{\sigma}\beta_6(age^2 \times time) + {}^{\sigma}\beta_7(age \times time^2)$$

$$\nu = {}^{\nu}\beta_1 age + {}^{\nu}\beta_2 age^2 + {}^{\nu}\beta_3 time + {}^{\nu}\beta_4 time^2$$
$$+ {}^{\nu}\beta_5(age \times time) + {}^{\nu}\beta_6(age^2 \times time) + {}^{\nu}\beta_7(age \times time^2)$$

$$\log \tau = {}^{\tau}\beta_1 age + {}^{\tau}\beta_2 age^2 + {}^{\tau}\beta_3 time + {}^{\tau}\beta_4 time^2$$
$$+ {}^{\tau}\beta_5(age \times time) + {}^{\tau}\beta_6(age^2 \times time) + {}^{\tau}\beta_7(age \times time^2)$$

where

age: the ordered variable for the 5-year age group

$time$: the continuous variable for time in year centered at the time of the last survey

${}^{i}\beta_{jk}$: the coefficient for the covariate j for the parameter $i = (\mu, \sigma, \nu, \tau)$

With the BCPE method, the median was given by the estimated parameter for the location parameter $\hat{\mu}$. The prevalence of overweight and obesity (BMI ≥ 25 kg/m^2) and obesity (BMI ≥ 30 kg/m^2) were given by $\left[1 - \hat{F}(25)\right]$ and $\left[1 - \hat{F}(30)\right]$, respectively, where $\hat{F}(BMI)$ is the estimated cumulative density function for BMI that was assumed to follow a BCPE distribution.

<u>Direct method with age group as a nominal variable</u>

$$BMI_{median} = \sum_k [\beta_{1k} age_k + \beta_{3k}(age_k \times time) + \beta_{4k}(age_k \times time^2)]$$

$$\text{logit}\,[Pr(BMI \geq 25)] = \sum_k [\beta_{1k} age_k + \beta_{3k}(age_k \times time) + \beta_{4k}(age_k \times time^2)]$$

$$\text{logit}\,[Pr(BMI \geq 30)] = \sum_k [\beta_{1k} age_k + \beta_{3k}(age_k \times time) + \beta_{4k}(age_k \times time^2)]$$

where

age_k: the indicator variable for the 5-year age group k

$time$: the continuous variable for time in year, centered at the time of the last survey

β_{jk}: the coefficient for the age category k of the covariate j

<u>Direct model with age group as an ordered variable</u>

$$\begin{aligned} BMI_{median} &= \beta_1 age + \beta_2 age^2 + \beta_3 time + \beta_4 time^2 \\ &\quad + \beta_5(age \times time) + \beta_6(age^2 \times time) + \beta_7(age \times time^2) \\ \text{logit}\,[Pr(BMI \geq 25)] &= \beta_1 age + \beta_2 age^2 + \beta_3 time + \beta_4 time^2 \\ &\quad + \beta_5(age \times time) + \beta_6(age^2 \times time) + \beta_7(age \times time^2) \\ \text{logit}\,[Pr(BMI \geq 30)] &= \beta_1 age + \beta_2 age^2 + \beta_3 time + \beta_4 time^2 \\ &\quad + \beta_5(age \times time) + \beta_6(age^2 \times time) + \beta_7(age \times time^2) \end{aligned}$$

where

age: the ordered variable for the 5-year age group

$time$: the continuous variable for time in year centered at the time of the last survey

β_j: the coefficient for the covariate j

Predicted curves were plotted against actually observed 3 obesity indicators for each age group in order to assess goodness-of-fit of the models visually. Then, in order to assess the prediction performance quantitatively, absolute prediction errors were calculated. An absolute prediction error was defined as an absolute difference between predicted and observed values for each age group. A country average was calculated as a weighted mean of the absolute prediction errors across all the age groups. The weight used was n_k/n_{total}, where n_k is the number of observations in age group k, and n_{total} is the total number of observations that were used to construct the prediction model.

Uncertainty about error estimates were presented using 95% bootstrap percentile confidence intervals given from 2,000 bootstrap samples as described previously (see 2.2.6 Analysis methods). A stratum, from which sampling units were sampled, was defined as a combination of the administrative subunit of the country and the urban-rural area category. One exception was the Mexican ENN 1988 survey, for which a stratum was defined as a combination of region (higher level than state, 4 regions in total) and the urban-rural area category. For each bootstrapped sample, sampling weights were calibrated so that sex- and age-specific sums of sampling weights are equal to those of the original data. All the prediction models converged with all the bootstrap samples. Sampling weights were included in all the analyses.

4.3 Results

4.3.1 Study population

The distributions of characteristics of the study populations (Table 4-2) were similar in terms of age structure. The proportions of people who resided in urban areas appeared to increase during the study period in Colombia and Peru. Population educational level seemed to increase in all the 3 countries, which is consistent with the results in the previous chapters.

Table 4-2 Demographic characteristics of the study populations by survey*

	Mexico				Colombia†				Peru				
	1988	2000	2006	2012	2005	2010	2005	2010	2000	2005-7	2008	2009	2014
Number of observations													
Number of records	10,744	20,079	14,044	14,601	2,828	2,722	8,720	10,376	20,168	9,176	11,767	18,025	19,585
Weighted counts (000)	14,149	20,461	23,619	26,154	2,693	2,715	2,749	2,736	5,208	5,790	5,969	6,096	6,614
Age													
20-29	43%	44%	36%	35%	63%	60%	62%	62%	40%	37%	36%	36%	35%
30-39	33%	34%	36%	36%	37%	41%	38%	38%	34%	35%	33%	35%	34%
40-49	24%	22%	29%	29%	--	--	--	--	27%	29%	31%	29%	31%
Type of residence													
Urban	--‡	--‡	77%	79%	69%	73%	72%	74%	70%	71%	73%	75%	78%
Rural	--‡	--‡	23%	21%	31%	27%	28%	26%	30%	29%	27%	25%	22%
Educational attainment													
Primary	--	--	44%	30%	43%	40%	32%	25%	37%	32%	32%	30%	24%
Secondary	--	--	43%	52%	49%	49%	52%	54%	38%	35%	36%	39%	40%
Higher	--	--	12%	18%	8%	11%	16%	21%	26%	33%	32%	31%	36%
Missing	--	--	0%	0%	--	--	--	--	--	--	--	--	--

* All the numbers were calculated using sampling weights except the number of records.

† Observations of non-pregnant women aged 20-39 with a birth history within 5 years from the survey.

‡ Definition of urban and rural area were different from those used in 2006-2012, and figures are not shown since they were not directly comparable.

4.3.2 Fit of predicted curves

On average, the models in which the age group variable was handled as an ordered variable (referred as "models with ordered age") performed better than models in which the age group variable was handled as a nominal variable (referred as "models with nominal age") in all the 3 obesity indicators (Figures 4-1a, -1b, -1c). For example, the predicted curves for obesity prevalence (BMI $\geq$ 30 kg/m^2) in Peru for the age groups 25-29 and 40-44 years, the BCPE and direct models with ordered age were closer to the observed values as compared to their counterpart models with nominal age (Figure 4-1c, bottom). Similar patterns were seen for overweight and obesity prevalence (BMI $\geq$ 25 kg/m^2) in Mexico in the age groups 20-24 and 45-49 years (Figure 4-1b).

4.3.3 Prediction errors

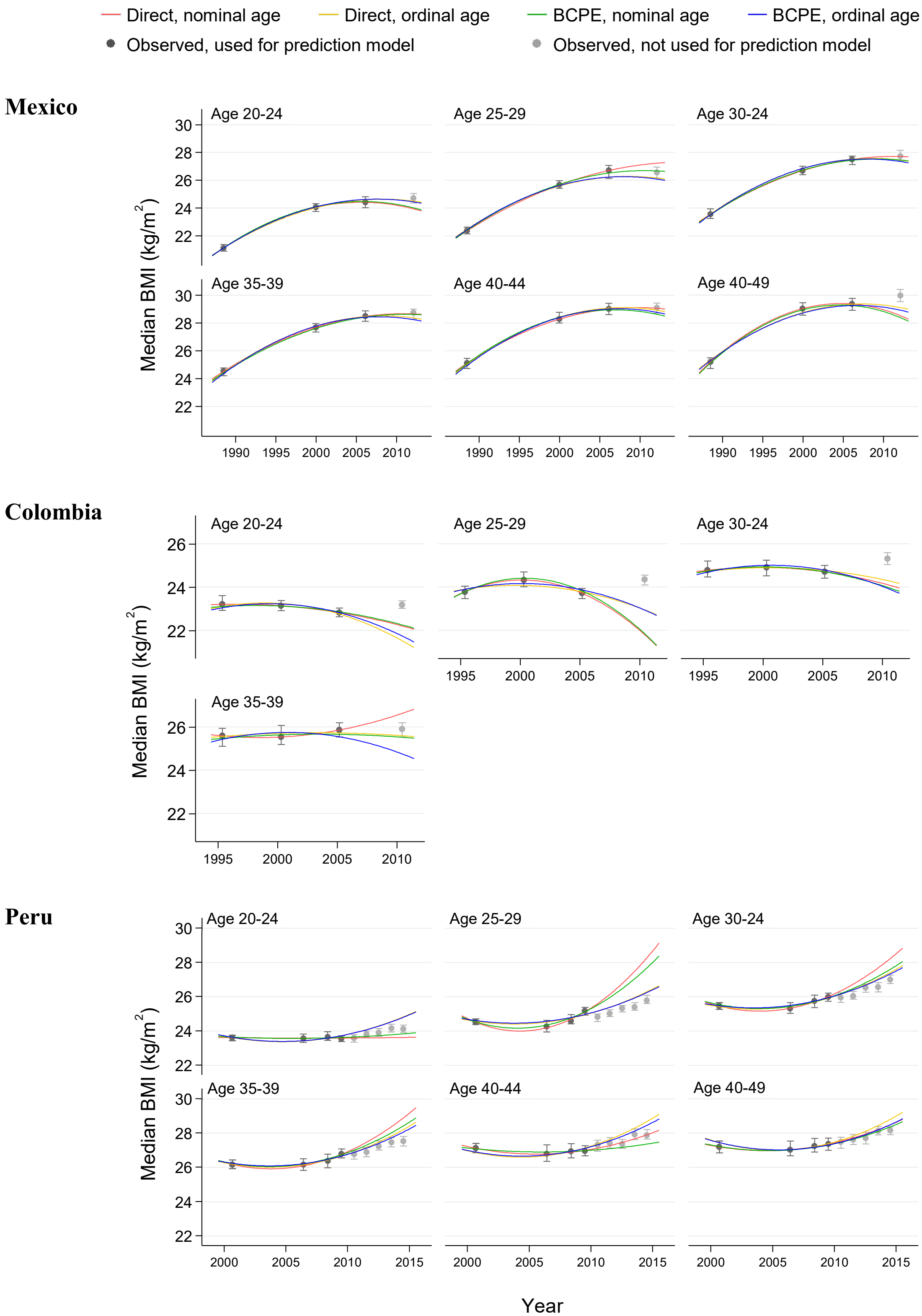
The BCPE model with ordered age yielded the best prediction performance as compared to the other 3 models in predicting obesity prevalence (BMI $\geq$ 30 kg/m^2) (Figure 4-2). Average absolute prediction errors across all age groups were 4.2 percentage points (pp) (95% percentile confidence interval: 1.9, 7.6), 2.5 pp (1.2, 6.1), and 1.7 pp (1.0, 9.2), with the data from Mexico, Colombia, and Peru, respectively. Errors of the direct model with ordered age were 6.2 pp (2.8, 9.8), 5.9 pp (3.0, 9.6), 4.7 pp (1.4, 14.8), respectively (Appendix G for age-specific results). Superiority in prediction by the BCPE model with ordered age was weak or none for overweight and obesity prevalence (BMI $\geq$ 25 kg/m^2) and median BMI.

4.4 Discussion

4.4.1 Main findings

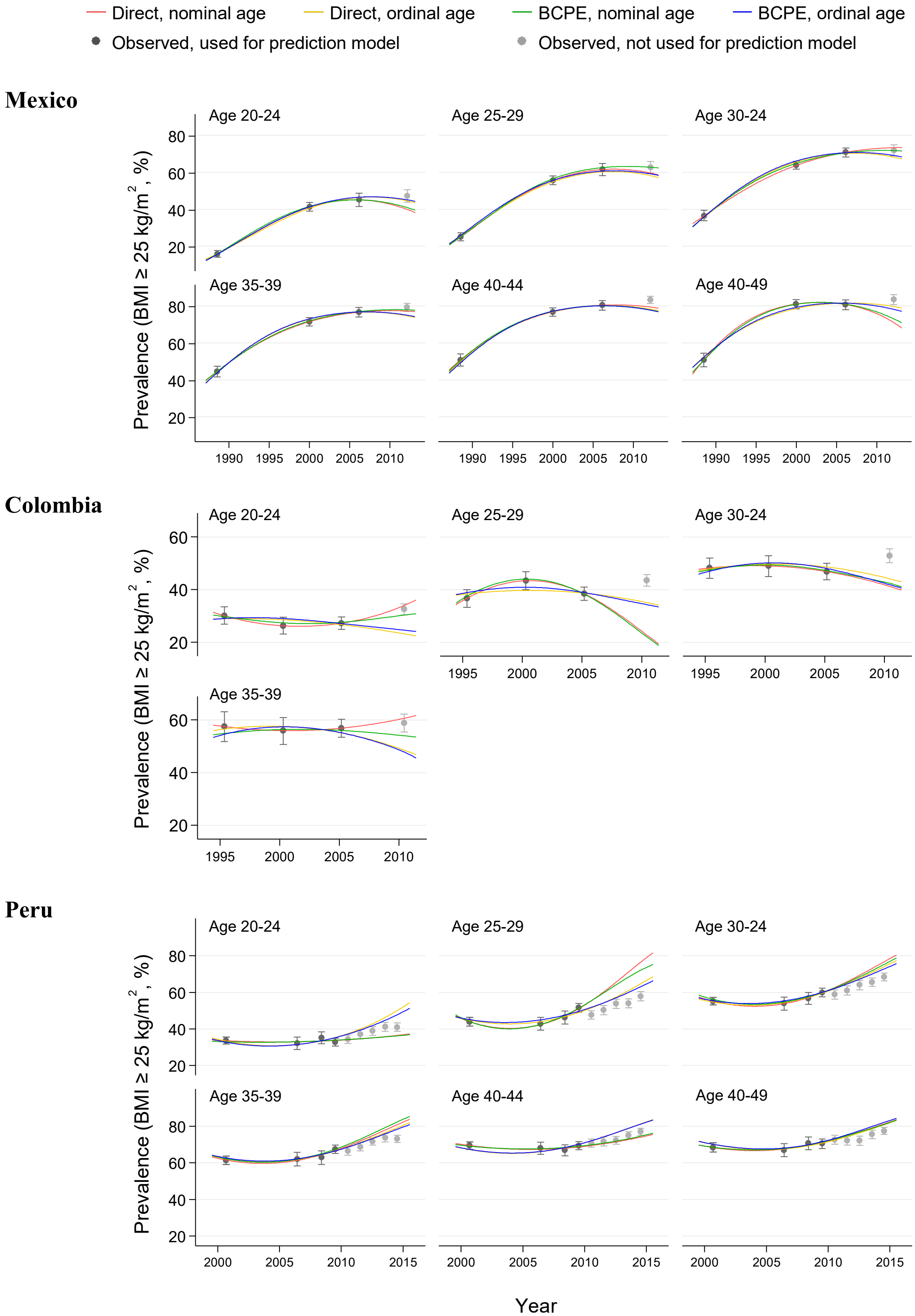
The BCPE method did better than the direct method in predicting the upper end of the BMI distributions (i.e., obesity prevalence, BMI $\geq$ 30 kg/m^2) but not much or no better for the other indicators (overweight and obesity prevalence, BMI $\geq$ 25 kg/m^2, and median BMI). Regarding the differences observed between BCPE and direct models, it might be possible that indicators closer to the center of

Figure 4-1a Predicted curves for median BMI



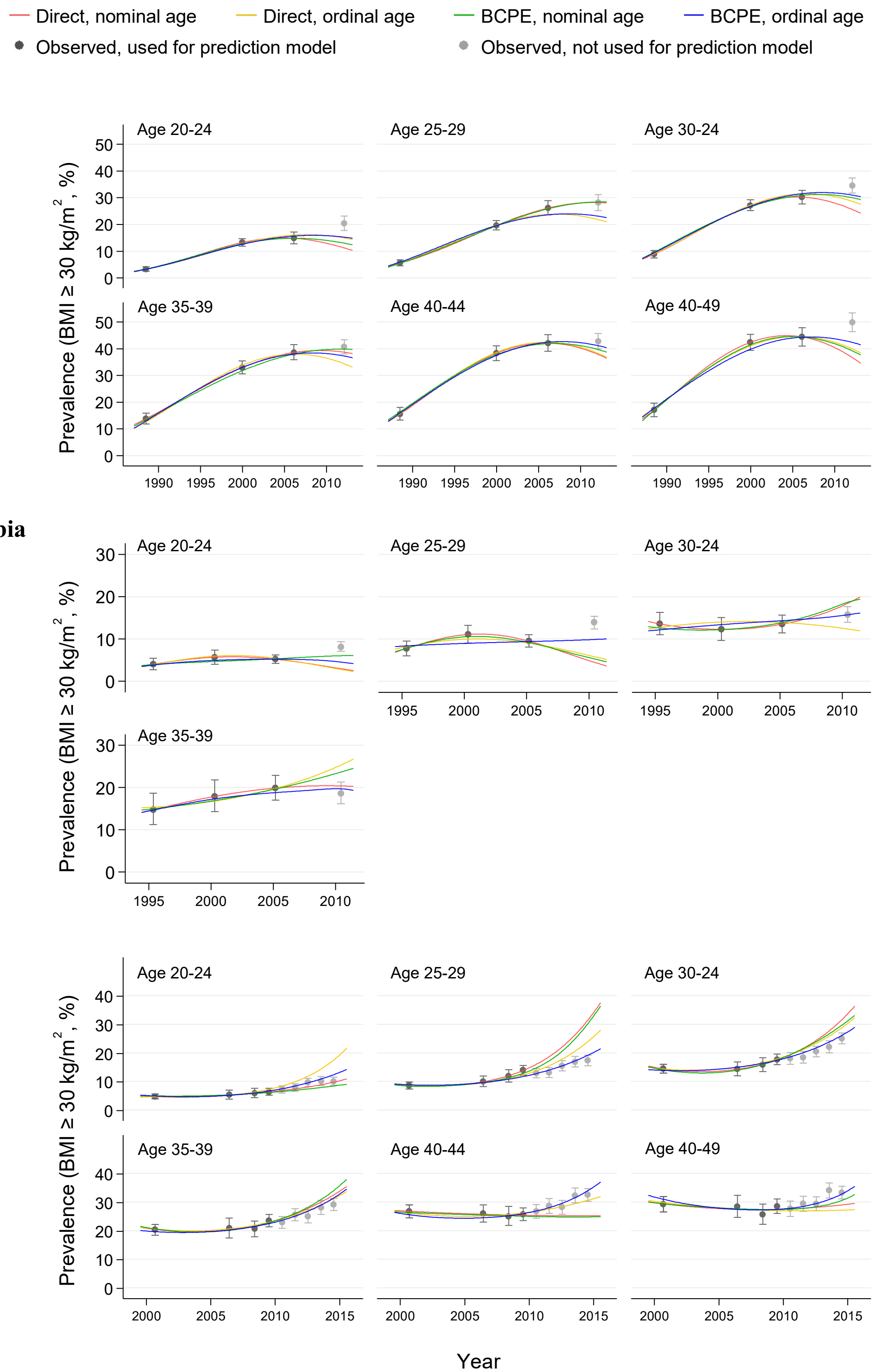
Direct, direct method; BCPE, BCPE method; nominal age, 5-year age group included as a nominal variable; ordinal age, 5-year age group included as an ordinal variable.

Figure 4-1b Predicted curves for overweight and obesity (BMI ≥ 25 kg/m^2) prevalence



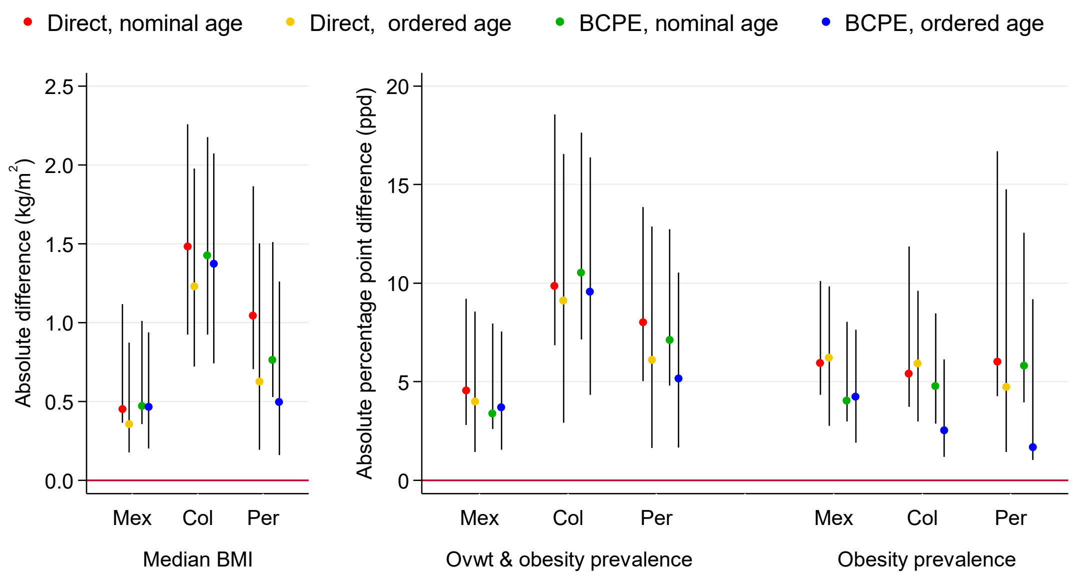
Direct, direct method; BCPE, BCPE method; nominal age, 5-year age group included as a nominal variable; ordinal age, 5-year age group included as an ordinal variable.

Figure 4-1c Predicted curves for obesity (BMI ≥ 30 kg/m^2) prevalence



Direct, direct method; BCPE, BCPE method; nominal age, 5-year age group included as a nominal variable; ordinal age, 5-year age group included as an ordinal variable.

Figure 4-2 Average absolute prediction errors



Indicator	Estimation model	Mexico		Colombia		Peru	
		Abs. pred. error	95% CI	Abs. pred. error	95% CI	Abs. pred. error	95% CI
Median BMI							
	Direct, nominal age	0.4	(0.4, 1.1)	1.5	(0.9, 2.3)	1.0	(0.7, 1.9)
	Direct, ordinal age	0.4	(0.2, 0.9)	1.2	(0.7, 2.0)	0.6	(0.2, 1.5)
	BCPE, nominal age	0.5	(0.4, 1.0)	1.4	(0.9, 2.2)	0.8	(0.5, 1.5)
	BCPE, ordinal age	0.5	(0.2, 0.9)	1.4	(0.7, 2.1)	0.5	(0.2, 1.3)
Overweight and obesity prevalence (BMI ≥ 25 kg/m^2)							
	Direct, nominal age	4.5	(2.8, 9.2)	9.9	(6.8, 18.6)	8.0	(5.0, 13.8)
	Direct, ordinal age	4.0	(1.4, 8.6)	9.1	(2.9, 16.6)	6.1	(1.6, 12.9)
	BCPE, nominal age	3.4	(2.6, 8.0)	10.5	(7.1, 17.6)	7.1	(4.8, 12.7)
	BCPE, ordinal age	3.7	(1.5, 7.5)	9.6	(4.3, 16.4)	5.2	(1.7, 10.5)
Obesity prevalence (BMI ≥ 30 kg/m^2)							
	Direct, nominal age	5.9	(4.3, 10.1)	5.4	(3.7, 11.9)	6.0	(4.3, 16.7)
	Direct, ordinal age	6.2	(2.8, 9.8)	5.9	(3.0, 9.6)	4.7	(1.4, 14.8)
	BCPE, nominal age	4.0	(3.0, 8.0)	4.8	(2.9, 8.5)	5.8	(3.9, 12.6)
	BCPE, ordinal age	4.2	(1.9, 7.6)	2.5	(1.2, 6.1)	1.7	(1.0, 9.2)

Mex: Mexico, Col: Colombia, Per: Peru, Obwt: overweight
Abs. pred. error: absolute prediction error
Direct, direct method; BCPE, BCPE method; nominal age, 5-year age group included as a nominal variable; ordinal age, 5-year age group included as an ordinal variable.

the distribution (e.g., median BMI and BMI ≥ 25 kg/m^2) change relatively steadily over time, but the extreme end of the distribution (e.g., BMI ≥ 30 kg/m^2) may not change in a simple manner. For this, modeling the entire BMI distributions considering their 4 components – location, scale, skewness, and kurtosis – might work better than modeling one aspect of the distribution.

The models with ordered age seemed more advantageous than those with nominal age by buffering outlying estimates in some age groups. For example, all obesity indicators estimated for Peruvian women aged 25-29 years had higher estimates in 2009 as compared to other adjacent age groups. These points drove the prediction curves upwards when the age group variable was handled as nominal (Figure 4-1a, -1b, -1c, bottom). By handling the age group variable as ordered, influence of such outlying estimates seemed to be reduced by borrowing trends in adjacent groups. Continuous age could have been used instead of 5-year age group; however, the use of age group was more appropriate in assessing the prediction performance of the models across different ages.

How to model time in the prediction model is debatable (Majer et al. 2013). Our study used a quadratic term of time, since the observed time trends were not linear (Figure 4-1a to 1c). The use of polynomial terms provides unstable estimates in general when the extrapolation point is further away from the data with which the prediction model is constructed. However, having only 3 time points, the use of a quadratic term was about the only option to handle observed non-linear time trends.

4.4.2 Strengths and limitations

This study, according to our best knowledge, is the first attempt to apply the BCPE-GAMLSS technique to predict future BMI distributions using survey data collected infrequently. Majer et al. used a prediction model in which the entire BMI distribution was predicted assuming that BMI follows a BCPE distribution. Their data were from Dutch Health Surveys that were conducted every year for 28 years, and time series technique was used for analysis. Our study is the second study that attempted to predict the entire BMI distributions but in a different manner. The data used in our study were collected about every 5 years or with longer intervals, and we did not have many time points of data. Our study was also unique in a sense that survey design (i.e., both sampling weight and clustering) was incorporated in the prediction.

There are several advantages in using this BCPE-GAMLSS approach as Majer et al. (2014) discussed in their report. The predicted prevalence of BMI categories always sums up to 1 since the prevalence is calculated as an area under the density function. This approach is also convenient since any quantity that depends on the distribution, such as the median BMI, and overweight and obesity prevalence, can be derived easily without fitting different models for each indicator. Although a clear difference between the BCPE and direct methods could not be demonstrated, it should be advantageous to be able to predict BMI distributions in a more precise manner. Lastly, the presented method can be applicable to countries where health surveys are conducted less frequently, which is mostly the case in low- and middle-income countries.

There are several limitations of our study. The study was based on data of women aged 20-49 years from the 3 countries. More countries should be added in order to assess the predictive performance of the BCPE model rigorously. Hence, our study is more indicative than definitive. Expanding the study population to men and postmenopausal women, which could not be done due to a lack of data at multiple time points, would provide more robust assessment of the models as well. However, the inclusion of a range of age groups from 20 to 49 years, at which patterns of change in BMI distributions differ most, may have captured major patterns of changes we could expect in other subpopulations.

Obviously, inclusion of more time points should result in more accurate prediction. However, this would be a challenge for most of the low- and middle income countries. A continuous survey, as Peru actually has implemented, may have an advantage in this regard. With more time points, the use of other smoothing techniques could become a possible option. Another way could be to borrow data from similar countries as implemented for the global estimate of obesity indicators (NCD Risk Factor Collaboration 2016) although this would require access to multiple datasets and high statistical capacities.

Chapter 5 Conclusions and directions for future research

BMI distributions and their rates of change differed by age, sex, place of residence (urban-rural), and socioeconomic factors in Mexico, Colombia, and Peru. While women had higher BMI in 2010 across ages, men generally experienced greater increases in BMI between 2005 and 2010. The young- to middle-age population tended to increase their BMI faster than other age groups in both sexes. Whereas the highest overweight and obesity prevalence was observed in Mexico in 2010, the highest rates of change in their prevalence was observed in Peru between 2005 and 2010. Opposite associations of education with BMI were observed between men and women – a positive association in men and a negative association in women. Household wealth was positively associated with BMI more clearly in men than women in all 3 countries. Lower wealth was associated with higher rates of change in BMI distributions in women.

These observations have some public health implications. We observed sex differences in patterns of BMI distributions and their changes by age, geographic and socioeconomic factors, which imply that the results from previous studies that used data of adult women are not necessarily generalizable to adult men. Differences across subpopulations within a country imply the need for tailoring obesity-related health policies and programs, if and where appropriate, so that they can reach the right target subpopulations with the right objectives (i.e., prevention of overweight and obesity vs. prevention of obesity-related complications and diseases). Effective approaches may differ by characteristics of target populations and by the nature of problems. It was also noticed there were some “catching-up” patterns among populations that had lower overweight and obesity prevalence than their counterpart groups as mentioned above – higher gains in BMI among the young to middle-age adults (vs. older adults), men (vs. women), Peruvian adults (vs. Mexican and Colombian adults), and women in the lower wealth groups (vs. those in higher wealth groups).

The proportion of adult populations with high BMI continued increasing among the 3 countries. This implies potential further increases of obesity-related diseases, such as cardiovascular diseases and diabetes, which will be accelerated by prolonged life expectancy. Since people with such diseases suffer from chronic conditions that require continuous control and treatment, financial burdens for the individuals and the health systems could increase enormously. Hence, a review of health care services and health systems might be also necessary. In addition, these chronic diseases cannot be cured instantly, and for the prevention and treatment of obesity and obesity-related diseases, both individual- and population/policy-level interventions are essential. Among those, policy interventions (e.g., taxation on sugar sweetened beverages, control of trans and saturated fat contents in food products, regulation of food advertisement for children) require dialogues and coordination with food industries, which involve great effort and time. Therefore, early mitigation of overweight and obesity problems and forward-thinking preparation for future burdens of such diseases would be indispensable.

Several components of the study may require future research. First, the study might be expanded to include adolescents, and children possibly, in order to have integrated views and better diagnosis of BMI distribution patterns and their changes across ages within each country. Although our study focused on adult populations, the problems of overweight and obesity start from childhood. For example, our study showed that more than 48% of young men and women aged 20-24 years were already overweight or obese as of 2010 in Mexico. Second, it would be indispensable to conduct further analyses, quantitative and qualitative investigations in order to understand possible causes of disparities we observed and to reformulate relevant health policies and programs. Our study provided descriptive information about BMI distributions and their change over time among subpopulations within each country. Regional differences are also expected within a country. Third, studies with other countries could be added to capture a wider view of obesity situations in Latin America. Our study covered 3 countries that are at different developmental stages among upper-middle-income countries in Latin America; however, countries are obviously different, and our study did not include low-, lower-middle-, and high-income countries in the region. Argentina, Chile, and Brazil conduct health surveys including both men and women. Other low- and middle-income countries implement DHS or other health surveys although many of them may be limited to women of reproductive age.

The method in which BMI is assumed to follow a BCPE distribution seemed to work well for both estimation and prediction of BMI. The BCPE distribution fitted well with the actual BMI distributions of the 3 countries. Comparison of BMI distributions across subpopulations and/or by time enabled us to understand dynamics of BMI distributions visually and easily – for example, whether the entire population experiences increases in BMI or only those with already higher BMI gain more. The BCPE prediction method yielded better performance in predicting obesity prevalence (BMI $\geq$ 30 kg/m^2) than the method in which the obesity indicators were modeled directly. These could be a result of modeling BMI distributions in a more precise manner considering not only the location and scale of distributions but also their shapes.

It should be also mentioned that our study is the first attempt to apply the method, in which BMI distributions are modeled for estimation and prediction, considering both sampling weight and clustering. Hence, the method presented in our study should be applicable to other studies with health survey data where sampling weight and clustering need to be incorporated. On the other hand, we were not able to assess the goodness-of-fit of the models in a rigorous manner, and this would be an area where further research would be necessary.

The data used for Mexico and those used for Peru in the first 2 objectives, were not from the DHS. While datasets and questionnaires were available on the website of the corresponding institutions, descriptions of datasets were not always available, and survey designs were not fully explained in the final reports or other sources. Data cleaning involved a great amount of effort. One of the superior features of the DHS is the use of standardized questionnaires, reporting formats, and data organization, which have been revised with accumulated knowledge and experience with numerous participating countries. Some of the DHS's approaches could be revisited when survey data management is reviewed in each country.

Appendix A Previous multi-country studies on adult obesity and its determinants in low- and middle-income countries

Table A Summary of previous multi-country studies on adult obesity and its determinants by category

Studies with data 2 time points or more that measured changes in variables of interest over time

Authors and year	Outcome variable	Assessed factors of interest	Analysis methods	Major findings	Data source	Study population
Lopez-Arana et al. 2014	Overweight and obesity (BMI ≥ 25)	Occupation (adjusted for education, wealth and other individual characteristics)	Binomial and Poisson regression	[1] SES In all analyzed 4 regions, women working in agriculture had consistently lower overweight and obesity prevalence, while women from professional, technical, managerial, and clerical classes had higher prevalence. [2] SES In Latin America and the Caribbean, South and Southeast Asia, and Sub-Saharan Africa, women working in agriculture and production experienced the largest increases in overweight and obesity while women in higher occupational classes experienced smaller increases.	DHS at 2 time points from 33 countries conducted in 1992-2009	Non-pregnant women aged 25-49
Mamun and Finlay 2014	Overweight and obesity (BMI ≥ 25), underweight, BMI < 18.5)	Education, household wealth, urban residence, and other individual and household characteristics	Multinomial logistic regression	[1] Overweight / obesity prevalence significantly increased in 27 countries (out of 36). [2] Annual increase of overweight and obesity prevalence was higher than annual decline of underweight prevalence (6.4% vs. 3.3%). [3] SES Whereas higher socio-demographic factors and urban residence were associated with shifting of underweight towards overweight, risk of highly educated, wealthy, and urban women being overweight was weakening.	DHS at 2 time points from 36 countries conducted in 1991-2008	Non-pregnant women 20-49
Neuman et al. 2014	Population mean BMI, overweight and obesity (BMI ≥ 25), underweight (BMI < 18.5)	GDP per capita, FDI, tariff, household wealth, urban residence (adjusted for other individual characteristics)	Multilevel linear regression, multinomial regression	[1] NatDev In unadjusted models: there was an insignificant positive association between annual change in GDP and change in BMI; no associations of FDI and tariff with the change in BMI were found. [2] NatDev, Geo, SES In adjusted models: GDP was associated positively with BMI, but its association was weaker in urban areas (vs. rural); GDP was associated positively with BMI except women in the highest wealth in whom the association was negative.	DHS at 2-3 time points from 38 countries conducted in 1991-2010	Non-pregnant women aged 15-49
Neuman et al. 2013	BMI, overweight and obesity (BMI ≥ 25), underweight, (BMI < 18.5)	Urban residence, household wealth (adjusted for age, education, marital status)	Linear regression, ordered logistic regression	[1] Geo Overweight increased more quickly in urban area in low-income countries, but it increased in rural area in upper-middle-income countries. [2] Geo Mean BMI was generally higher in urban area, but the association was attenuated after adjusted for socioeconomic status. [3] SES Individual- and household-level SES measures were independently and positively associated with BMI.	DHS at 2 time points from 38 countries conducted in 1991-2010	Non-pregnant women 15-49

Jones-Smith et al. 2012	Overweight and obesity (BMI ≥ 25)	Education, household wealth, GDP per capita	Difference and growth rate of prevalence, correlation analysis	[1] SES The highest wealth and education groups have the highest prevalence of overweight and obesity in the majority of countries. [2] SES Increases in overweight prevalence were greater in the lowest wealth group than the highest group in 11 countries (out of 39). [3] SES Increases in overweight prevalence were greater in the lowest education group than the highest group in 21 countries (out of 39). [4] SES, NatDev GDP was positively associated with growth rate of overweight and obesity prevalence. Such an association was not observed between GDP and education.	DHS at 2 time points from 39 countries conducted in 1991-2008.	Non-pregnant women 18-49 with birth history within 5 yrs from survey
Jones-Smith et al. 2011	Overweight and obesity (BMI ≥ 25)	household wealth; GDP per capita, change in GDP (age controlled by standardization)	Change in Slope Index of Inequality (SII), meta-regression of change in SII	[1] SES In 10 countries (out of 37), higher wealth was negatively associated with gains in overweight prevalence. [2] SES, NatDev GDP was positively associated with faster increase in overweight prevalence among the lower wealth groups.	DHS and other national surveys at 2 time points from 35 countries conducted in 1989-2007.	Non-pregnant women 18-49 with birth history within 5 yrs from survey
Neuman et al. 2011	BMI, overweight and obesity (BMI ≥ 25)	Household wealth, education (adjusted for age, education, place of residence)	Linear and Poisson regression	[1] SES The association between BMI and wealth was positive in 37 countries (out of 37) in the first round of surveys and in 36 countries in the second round of surveys. [2] SES In 6 countries (out of 37), increases in BMI between 2 surveys were higher among the lowest wealth group than the highest.	DHS at 2 time points from 37 countries conducted in 1991-2008.	Ever-married non-pregnant women aged 15-49

Cross-sectional studies and systematic reviews

Authors and year	Outcome variable	Assessed factors of interest	Analysis methods	Major findings	Data source	Study population
Goryakin et al. 2015	Overweight and obesity (BMI ≥ 25)	Globalization indices, GDP, HDI. Economic freedom score (adjusted for individual demographic characteristics)	Scatter plot with lowess curves, linear regression.	[1] NatDev Total, economic and social globalization indices were linearly associated with overweight and obesity prevalence. Political globalization index was linearly associated with overweight and obesity prevalence from a certain level of the index.	DHS from 56 countries conducted in 1991-2009. Multiple surveys included per country.	Non-pregnant women 15-49

Aitsi-Selmi et al. 2014	Obesity (BMI ≥ 30)	Education, household wealth (adjusted for age, parity, place of residence)	Logistic regression	[1] SES In 4 middle-income countries, wealth was positively associated with obesity among women with no/primary education, but it had no or negative association among those with higher education. [2] SES In 3 low-income countries, positive and independent association of wealth and education with obesity were observed.	Largest DHS from 7 countries conducted in 2005-10.	Non-pregnant women 15-49
Goryakin and Suhrcke 2014	Overweight and obesity (BMI ≥ 25)	GDP per capita, individual education, occupation, and other national and household characteristics.	Scatter plot with lowess curves, linear regression.	[1] SES The relationship between GDP per capita (PPP, 2005 values) and overweight and obesity prevalence was positive and peaked at about US$ 5,000-6,000. [2] SES, NatDev The association between education and overweight and obesity was positive in low income countries and negative in medium-income countries. [3] Geo, NatDev Urban residence was significantly related with higher overweight and obesity across the countries at all income level. [4] SES, NatDev Shifting patterns of employment from agriculture into services was significantly related with higher overweight and obese across the countries at all income level.	DHS from 56 countries conducted in 1991-2009. Multiple surveys included per country.	Non-pregnant women 15-49
Nandi et al. 2014	Population mean BMI, overweight and obesity (BMI ≥ 25), underweight (BMI < 18.5)	GDP per capita, %urban population, FDI, tariff, education, income, and other individual characteristics	Meta-regression, multinomial regression	[1] NatDev In the unadjusted model, a non-linear positive association of GDP with population mean BMI was observed. There was a positive association between %urban population and population mean BMI. [2] NatDev, SES, Geo In the adjusted model, a positive association of GDP with overweight and obesity prevalence was observed. Little association was observed for %urban population and education. There was a positive association of income in men and women with overweight and obesity in both urban and rural areas.	WHS from 40 countries conducted in 2002-2003. Self-reported weight and height. Multiple imputation for missing data.	Men and women aged 18+
Van Hook et al. 2013	Overweight and obesity (≥ 85^{th} BMI percentile for women < 20, BMI ≥ 25 for women 20+)	GNP, urban residence, education (adjusted for age, marital status, year of survey, country)	Logistic regression,	[1] NatDev, SES, Geo There were positive associations of GNI, education, and urban residence with overweight and obesity. [2] NatDev, SES, Geo There were positive associations between GNP and overweight and obesity except women with higher education in whom a negative association was seen. There were positive associations between GNI and overweight and obesity, but the association was less steep in urban areas than rural.	DHS and other national surveys at 2-3 time points from 33 countries conducted in 1990-2008.	Non-pregnant women aged 15-49

Corsi et al. 2012	BMI; overweight and obesity (BMI ≥ 25), underweight (BMI < 18.5)	PSU as neighborhood indicator (adjusted for age, education, household wealth, place of residence)	Multilevel linear regression, multinomial regression	[1] Geo Variabilities in BMI, overweight and underweight prevalence between neighborhood varied substantially across the counties. In countries with greater neighborhood variation, BMI might be influenced by local conditions more.	DHS from 57 countries conducted in 1994-2008.	Non-pregnant women aged 20-49
Dinsa et al. 2012	Obesity	SES (income, assets, education, occupation, etc.)	Systematic review	[1] SES, NatDev In low-income or low HDI countries, the association between socioeconomic status and obesity appeared to be positive in men and women. In middle-income or medium HDI countries, the association became mixed for men, and mainly negative for women.	Published studies in 2004-2010 for low- and middle-income countries.	Adult men and women
Egger, et al. 2012	Population mean BMI	GDP per capita	Linear regression with splines	[1] NatDev GDP (2007 values) was positively associated population mean BMI up to about USD 3,000.	Country-level estimated data of 175 low- to high-income countries.	Adult men and women
Fleischer et al. 2012	BMI	Education (adjusted for age, stratified by sex)	Linear regression for each country, then meta-analysis for pooled analysis	[1] SES, Geo Persons with higher education had a higher BMI in the least urban countries; the opposite pattern was seen in the most urban countries, especially among women.	WHS from 70 countries in 2002-3. Self-reported weight and height.	Men and women aged 18+
Pampel et al. 2012	BMI; proportion of 4 BMI categories	Sex, age, marital status, education, occupation, household wealth, urban residence; GDP per capita	Linear regression, multinomial logistic regression	[1] NatDev, SES In the single-term model, positive associations of wealth and GDP with BMI and overweight and obese were found in men and women. Positive associations of wealth with BMI and overweight and obese existed more among men than women. A negative association of education with BMI and overweight and obese existed in women, and a weak negative association existed in men. [2] SES, NatDev In the mode with interactions between SES variables (education, non-manual occupation wealth) and GDP, all SES variables had positive association at lower GDP, but this association attenuated and reversed at higher GDP in both men and women.	WHS from 67 countries conducted in 2002-3. Self-reported weight and height. Multiple imputation for missing data.	Men and women aged 18+

Subramanian et al. 2011	BMI; overweight and obesity (BMI ≥ 25)	Education, household wealth, place of residence, GDP per capita (adjusted for age)	Linear regression, logistic regression	[1] SES One average, an increase in wealth was associated with an increase in BMI and an increase in overweight. [2] SES 52 (out of 54) countries had positive associations between wealth and BMI and overweight. [3] Geo The positive association between wealth and BMI was stronger in urban areas than rural areas. [4] SES, NatDev Positive associations between GDP and BMI or overweight were observed, but the association was less among women in the highest wealth group.	DHS from 54 countries conducted in 1995-2008.	Women 15-49
McLaren 2007	Obesity and synonymous	SES (education, income, material possessions, occupation, etc.), HDI	Systematic review	[1] SES, NatDev Overall, lower SES was associated with higher body size in high-HDI countries; this association started attenuating and reversing as moved to lower HDI countries. [2] SES, NatDev Negative associations between SES and large body size were common among women in high-HDI countries; positive associations between SES among women in low- to middle-HDI countries; mixed associations among men in medium- to high-income countries.	333 published studies in 1988-2004, including reports from low- to high-income countries.	Men and women aged 18+
Ezzati, et al. 2005	Population mean BMI	GDP per capita (age controlled by standardization)	Scatter plots, lowess curves.	[1] NatDev BMI increased most rapidly until a GDP per capita of about I$ 5,000 (international dollar), and peaked at about I$ 17,000 for men and I$ 12,500 for women. [2] SES, Geo There was an inverse relationship between BMI and the food share of household expenditure; a positive relationship between BMI and proportion of urban population.	Published studies and reports from 69 low- to high-income countries.	Men and women aged 30+
Mendez et al. 2005	Overweight and obesity (BMI ≥ 25), underweight prevalence, (BMI < 18.5)	GNI, level of urbanization stratified by urban-rural area (age controlled by standardization)	Descriptive analysis, cubic spline curves, linear regression with log prevalence as outcome	[1] NatDev, Geo Country with high GNIs and high levels of urbanization had: high overweight and obesity prevalence; small urban-rural difference in overweight and obesity prevalence; and high ratio of overweight and obesity to underweight prevalence. [2] SES, Geo In more developed countries, overweight and obesity prevalence among low education group was high in both urban and rural areas.	DHS and other national survey from 36 countries conducted in 1992-2000.	Non-pregnant women aged 20-49
Monteiro et al. 2004a	Obesity (BMI ≥ 30)	GNP per capita, education (age controlled by standardization)	Scatter plot; logistic regression	[1] NatDev There was a curvilinear positive (steep slope first, then leveling off) association between GNP and obesity prevalence. [2] SES, NatDev In low-income countries, obesity tended to increase by education; in lower-middle-income countries, the pattern of association was mixed; in upper-middle-income countries, obesity decreased by education. [3] SES, NatDev At GNP per capita of US$ 2,500 (2001 values) and above, obesity prevalence was higher among women in the lowest education group than those in the highest.	DHS and other national survey from 37 countries conducted in 1992-2000.	Non-pregnant women aged 20-49

Monteiro et al. 2004b	Obesity	GNP per capita, SES (measured by income, education, or composite indicator)	Systematic review, linear regression	[1] SES, NatDev As GNP increases, obesity tended to shift to those with low SES in men and women. [2] SES, NatDev The shift of obesity to those with low SES occurred at earlier stage of economic development in women than men.	Published studies in 1989-2004; 14 studies conducted in 1982-2002.	Adult men and women
Martorell et al. 2000	Obesity (BMI ≥ 30)	GNI per capita, place of residence, education	Scatter plots	[1] NatDev Obesity increased until a GDP per capita of USD 1,500 (1992 values) and changed a little thereafter. [2] SES, NatDev In poorer countries, obesity was concentrated among urban and high educated women. In mode developed countries, obesity was more equally distributed.	DHS and other national surveys from 39 countries conducted in 1989-1996. (some countries contributed twice)	Non-pregnant women aged 15-49
Sobal and Stunkard 1989	Obesity	SES (measured by income, education, occupation, or composite indicator)	Systematic review	[1] SES, NatDev In developed countries, there was a strong inverse association among women, and inconsistent association among men. In developing countries, there was a strong direct association among men and women.	144 published studies from low- to high-income countries; studies conducted in 1933-1988	Adult men and women

SES findings related with socioeconomic status; Geo findings related with urban-rural residence; NatDev findings related with country development.

BMI, body mass index; DHS, Demographic and Health Survey; FDI, foreign direct investment; HDI, Human Development Index; GDP, gross domestic product; GNI, gross national product; PSU, primary sampling unit; SD, standard deviation; SES, socioeconomic status; WHS, World Health Survey

Appendix B Previous studies on BMI prediction and their methods

Table B Summary of previous studies on BMI prediction by category

Studies that directly predicted obesity indicators

Authors and year	Prediction method	Predicted parameters	Included covariates	Stratifiers	Data source
Mills 2009	Composition data analysis; response variable was proportions of each BMI category.	Proportion of each BMI category (not overweight, overweight, obese)	Year	Sex, age	UK HSE 1993-2005
Zaninotto, et al. 2009	Obesity prevalence was calculated for each year, by sex, age group, and social class. Then, linear, power, and exponential curves were fitted.	Obesity prevalence	Year	Sex, age, social class	UK HSE 1993-2004
Wang et al. 2008	Linear regression.	Overweight and obesity prevalence	Year	Sex, age, race	NHANES 1971-2004
	Obtained BMI percentiles for each survey time; fitted a linear regression for each percentile.	Mean BMI in each percentile	Year	None	
Wang et al. 2007	Linear regression on survey data.	Mean BMI	Age, year, their higher-order terms, their interaction	Sex, race	US NHANES 1971-2004
Ruhm 2007	Quantile regression for each percentile of BMI (1^{st} to 99^{th}). From the predicted mean BMI at each percentile, prevalence of obesity was linearly interpolated from adjacent percentiles.	Obesity prevalence	Age, race, year	Sex	US NHANES 1960-1999

Studies that predicted obesity indicators borrowing data of other countries

Authors and year	Prediction method	Predicted parameters	Included covariates	Stratifiers	Data source
NCD Risk Factor Collaboration 2016	Bayesian hierarchical model.	Mean BMI Prevalence of BMI categories	Age, year, national income, urbanicity, measures of national food availability, study level indicators (representativeness, area of coverage), mean years of education, measures for availability of food types	Sex	Published and unpublished population representative surveys and studies
Ng et al. 2014	Spatiotemporal regression model and Gaussian process regression.	Overweight and obesity prevalence	Age, year, region, annual total kilocalories consumed, latitude, urbanicity	Sex	Population representative surveys and studies
Stevens, et al. 2012	Bayesian hierarchical mode to estimate mean BMI; then, regression models were used to estimate overweight and obesity prevalence, including estimated mean BMI as a covariate.	Overweight and obesity prevalence	Estimated mean BMI, age, sex, year, income category	None	Published and unpublished health surveys and epidemiologic studies
Finucane, et al. 2011	Extracted data on mean BMI and/or overweight /obesity prevalence by sex and age, with indicators of sampling variability, survey population, and sampling strategy. Then, Bayesian hierarchical mode was applied.	Mean BMI	Age, country, year, national income, urbanization, national availability of multiple food types	Sex	Published and unpublished health surveys and epidemiologic studies

Studies that predicted obesity indicators by microsimulation

Authors and year	Prediction method	Predicted parameters	Included covariates	Stratifiers	Data source
Webber, et al. 2012	Same as Wang et al. 2011	Proportion of 3 BMI categories (not overweight, overweight, obese)	Year	Sex, age	Published and unpublished data and publications

Wang et al. 2011	Categorical regression with a constrain that the proportions of each BMI category sum up to 1; microsimulation with age-, sex- and year-specific probabilities obtained from the regression model. Assumed that individual's BMI ranking in the same-age cohort constant over time.	Proportion of 3 BMI categories (not overweight, overweight, obese)	Age, year	Sex	UK HSE 1993-2008 NHANES 1999-2008
Basu 2010	Probabilistic modeling using annual transitional probabilities between 5 BMI categories, and age-, sex-, race-, BMI category-specific mortality. The transitional probability was obtained by conditional ordered logit regression on longitudinal survey data.	Proportion of population in 5 BMI categories (normal, overweight, obese class 1-3)	BMI categories, race, sex, age, age^2, year, interaction terms between BMI categories and other covariates	Age categories (6-17, 18-85)	US Medical Expenditure Panel Survey 2001-2, 2004-5
McPherson et al. 2007	Categorical regression with a constrain that the proportions of each BMI category sum up to 1; microsimulation with category-specific probabilities obtained from the regression model.	Proportion of 5 BMI categories (underweight, appropriate, overweight, obese, morbidly obese)	Age, sex, year, ethnicity, social class, geographical region (not all of them at a time)	Country (US, UK)	UK HSE 1993-2004

Prediction of the entire BMI distribution

Authors and year	Prediction method	Predicted parameters	Included covariates	Stratifiers	Data source
Majer et al. 2012	Generalized additive models for location, scale and shape (GAMLSS) semi-parametric regression assuming that BMI follows a Box-Cox power exponential (BCPE) distribution. The parameters were estimated for each combination of sex, age and year; then the parameters were modeled and predicted using Lee-Carter model.	BMI distribution (4 parameters of the BCPE distribution)	Age, sex, year	None	Dutch Health Survey 1981-2008
Sperrin et al. 2014	Latent class analysis.	Per year BMI, latent classes	Age, year	Sex	UK HSE 1992-2010

NHANES: National Health Examination Survey, UK HSE: Health Surveys for England, UK

Appendix C Sampling designs of surveys used in the study

Table C Summary of sampling designs by survey

Mexico

Survey name and data col. period	Design	Definition of urban, rural area	Reference

National Health and Nutritional Survey (ENSANUT)
 	
ENSANUT 2006*:
Oct 2005 – Apr 2006
 	
ENSANUT 2012*:
Oct 2011 – Mar 2012
 	
* Dates of data collection are not available in the final reports and obtained from the datasets.
	- **PSU**: Basic geostatistical area (AGEB, Spanish abbreviation), defined as a conventional group of 20-80 blocks surrounded by streets in urban area or about 10,000 hectors of area surrounded by natural borders in rural area.
	- **Stratification**: In the 2006 survey, in each state, AGEBs were grouped into 6 strata, which were created as a combination of 3 types of localities (metropolitan, urban, rural) and 2 categories of the Oportunidades Program, a social protection program (incorporated or not). The latter category was added in order to evaluate the program. In the 2012 survey, in each state, AGEBs were grouped into 7 strata, which were created as a combination of 3 type of localities (same as the previous survey) and 2 categories of the social gap index (low or high) plus newly emerged localities in the recent census.
	- **Sampling**:
	  - AGEBs were sampled from each stratum within each state with probability proportional to size (PPS) of dwellings.
	  - In metropolitan and urban strata: within an AGEB, 6 blocks were sampled with PPS; from each block, 6 dwellings were sampled systematically (in the 2006 survey) or randomly (in the 2012 survey).
	  - In rural strata: In the 2006 survey, from each AGEB, secondary sampling units (SSUs), which consisted of one or more localities, were selected with PPS; from each SSU, 3 conglomerates of 12 dwellings were sampled systematically. In the 2012 survey, from each AGEB, 3 localities were sampled with a PPS; from each locality, one pseudo-block, which consisted of about 50 dwellings, was sampled randomly; and from each pseudo-block, one conglomerate of 12 dwellings was sampled.
	  - All households within the selected dwellings were eligible for the survey. From each household, one adult (in the 2006 survey) or at least one adult (in the 2012 survey) were sampled randomly from each household.
	- **Data collection**: Data were collected through household and individual questionnaires, anthropometry, and blood specimens. The health team collected information about household and household member characteristics, and health conditions and service use. Nutrition team made anthropometric measurement, measured blood pressure, collected blood specimens, and surveyed about nutrition-related behaviors, and participation to social and health protection programs.
		Metropolitan: Localities with a population of ≥100,000, in the state capital.
		
		Urban: Localities with a population of 2,500 – <100,000 that is not metropolitan.
		
		Rural: Localities with a population of <2,500.
		
		* In this study, metropolitan and urban localities were considered as urban.
			Sepúlveda-Amor, et al. 1990; Olaiz-Fernández, et al. 2006; Shamah-Levy, et al. 2007; Gutiérrez, et al. 2013.

National Health Survey (ENSA) Sep 1999	- **PSU**: Municipalities. - **Stratification**: A stratum consisted of a combination of state and type of municipalities (urban or rural). - **Sampling**: - Number of dwellings was assigned to each strata (urban or rural) per state.	**Urban**: localities with a population of ≥15,000.	Valdespino et al. 2003; Olaiz et al. 2003.

– Mar 2000	- From each stratum, 14 municipalities were sampled with replacement with probability proportional to size (PPS) of dwellings. - From each selected municipality, 5 AGEBs were sampled with PPS. - From each selected AGEB, 3 blocks were sampled with equal probability. - From each selected block, 7 dwellings were sampled with equal probability. - All households within the selected dwellings were eligible for the survey. - From each selected household, one child, one adolescent, and one adult were sampled randomly from each household. - **Data collection**: Data were collected through household and individual questionnaires, anthropometry, blood and urine specimens.	**Rural**: localities with a population of <15,000.	
National Nutritional Survey (ENN) Apr – Sep 1988	- **PSU**: Sub-division of municipalities, which consisted of one or more AGEBs, with a minimum of 640 dwellings. - **Stratification**: In each state, municipalities were grouped in 3-12 stratus by their socioeconomic status. - **Sampling**: - PSUs were sampled from each stratum with probability proportional to size (PPS) of populations. - In urban area, within a selected PSU, 8 secondary sampling units (SSUs), which consisted of about 40 dwellings, were sampled with PPS. In rural area: within a selected PSU, 4 SSUs were sampled with PPS. - In urban area, from each selected SSU, 5 dwellings were sampled systematically. In rural area, from each selected SSU, 10 dwellings were sampled. - All children under 5 and women of reproductive age (12-49) were interviewed. - **Data collection**: Data were collected through household and individual questionnaires, anthropometry, and blood specimens.	**Urban**: municipalities where more than half of the population live in localities with a population of >15,000 **Rural**: municipalities where more than half of the population live in communities with a population of <15,000	Sepúlveda-Amor et al. 1990.

Table C Summary of sampling design by survey (continued)

Colombia

Survey name and data col. period	Design	Definition of urban, rural area	Reference

Demographic and Health Survey / Nutritional Situation National Household Surveys (ENDS / ENSIN)
 	
ENDS / ENSIN 2005:
Oct 2004 – June 2005
 	
ENDS / ENSIN 2010:
Nov 2009 – Nov 2010
	- **PSU**: One municipality or a group of municipalities (if with a population of 7,000).
	- **Stratification**: PSUs were grouped into strata within each department by similarities of their characteristics, which included population size in municipal capital, urban-rural ratio, living condition index, geographic neighborhood.
	- **Sampling**:
	  - From each stratum within a department, PSUs were sampled. PSUs, whose size was similar to the average stratum size, constituted PSUs with a selection probability of 1. Other PSUs were sampled with probability proportional to size (PPS) of population.
	  - Within each selected PSU, urban blocks were sampled systematically in municipal capital(s); rural sections were sampled systematically in the rest of the PSU.
	  - From each selected secondary sampling unit, segments were sampled with probability proportional to size (PPS) separately for urban and rural area. A segment consisted of an average of 10 dwellings.
	  - All households within the selected dwellings were eligible for the survey. From each household, all household members aged between 0 and 64 were measured for anthropometry. Subsamples were taken for evaluation of dietary patterns, physical activities, determinants of nutritional status, and biochemical indicators.
	- **Data collection**: Data were collected through household and individual questionnaires, anthropometric measurement, blood specimens. In the 2005 survey, ENDS teams collected information about household and household member characteristics, maternal and child health (including anthropometry) whereas ENSIN teams surveyed about nutritional health. In the 2010 survey, one unified teams collected information.
		Urban: Capital area within a municipality.
		
		Rural: Other areas within a municipality.
			Proforma & Macro International 2005a; ICBF 2006; ICBF 2011; Profamilia & ICF Macro 2011a.

Demographic and Health Survey (ENDS)
 	
ENDS 1995:
Mar – Jun 1995
 	
ENDS 2000:
Feb – Jun 2000
	- **PSU**: One municipality or a group of municipalities (if with a population of 7,000).
	- **Stratification**: PSUs were grouped into strata within 14 sub-regions (higher level than department) by similarities of their characteristics, which were total population, population in the municipal capital, proportion of rural population, department, and socioeconomic characteristics.
	- **Sampling**:
	  - From each stratum within a sub-region, PSUs were sampled. PSUs whose size was similar to the average stratum size constituted a stratum with selection probability of 1. Other PSUs were sampled using "controlled selection" technique with variables for additional stratification that included department and social development index.
	  - From each selected PSU, areas were sampled separately from municipal capitals (urban) and the rest (rural). An area consisted of an average of 60 dwellings or 6 segments, and one segment consisted of an average of
		Urban: Capital area within a municipality.
		
		Rural: Other areas within a municipality.
			Profamilia & Macro International 1995a, 2000a.

	10 dwellings. The number of segments to be selected for urban and rural was determined by population ratio between these two area. - From each selected area, segments were selected randomly. - All households within the selected dwellings were eligible for the survey. From each household, all children under 5 years of age and women of reproductive age were surveyed. - **Data collection**: Data were collected through household and individual questionnaires, and anthropometric measurement.		

ICBF, Instituto Colombiano de Bienestar Familiar

Table C Summary of sampling design by survey (continued)

Peru

Survey name and data col. period	Design	Definition of urban, rural area	Reference

National Household Survey, Module for Monitoring of Nutritional Indicators (ENAHO-MONIN)
 	
ENAHO-MONIN 2007-8: Nov 2007 – Mar 2008
 	
ENAHO-MONIN 2012-3: Jul 2012 – Sep 2013.
	- **PSU**: In urban area, PSUs were localities with a population of ≥2,000; in rural area, localities with a population 500 - <2,000, or Census Rural Areas with an average of 100 dwellings.
	- **Stratification**: PSUs were grouped into strata that were defined by combinations of natural regions, department, and urban-rural area.
	- **Sampling**: ENAHO is a continuous survey starting from 2003, in which some selected households were followed up longitudinally. However, there were no households that were surveyed by ENAHO-MONIN in both 2007-8 and 2012-3, which were used for this study.
	  - In urban area, from each PSU, secondary sampling units (SSUs, conglomerates with an average of 120 dwellings) were sampled with probability proportional to size (PPS) of dwellings. Then, dwellings were sampled systematically from selected SSUs.
	  - In rural area, if a locality was with a population 500 - <2,000, the sampling method was the same as that for urban area. If the locality is a Census Rural Area, dwellings were sampled systematically.
	  - In selected dwellings, all household members were measured for anthropometry.
	- **Data collection**: Main ENAHO collected information about household and household member characteristics that included household assets, access to basic services (e.g., water, sewage, electricity), demographic characteristics, educational, health, employment, income and expenditure, participation to social services through questionnaires. The following MONIN collected anthropometric measurements from all household members, and blood specimen from selected members.
		Urban: Localities with a population of ≥2,000.
		
		Rural: Localities with a population of <2,000.
		
		* The datasets did not have variables with this definition. Then, in this study, the following definitions were used.
		
		Urban: Localities with >400 dwellings.
		
		Rural: Localities with ≤400 dwellings or those categorized as Census Rural Areas.
			INEI 2012a, 2012b, 2013[a], 2014c; INEI & CENAN 2010a, 2012, 2014; DEVAN, 2015a.

Demographic and Health Survey (ENDSA)
 	
ENDSA 2009-2014: conducted continuously.
	- **PSU**: In urban areas, PSUs were conglomerates, each of which was consisted of one to several blocks with an average of 120 dwellings. In rural areas, PSUs were defined in the same way as the urban areas, or Census Rural Areas, which consisted of one or more populated centers with an average of 120 dwellings.
	- **Stratification**: PSUs were grouped into strata that were defined by combinations of department, and urban-semirural-rural area. Within each department, number of PSUs to be sampled were distributed according to the size of population of each stratum.
	- **Sampling**: ENDSA started to apply the continuous survey starting from 2003. ENDSA has gone through 3 cycles of continuous surveys (1st: 2003-2008, 2nd: 2009-2011, 3rd: 2012-2014). ENDSA 2000 was an independent survey.
	  - From each stratum, PSUs were sampled with probability proportional to size (PPS) of dwellings.
	  - Selected PSUs were randomly distributed, but maintaining urban-semirural-rural ratios within department, to 4-5 survey periods within the survey cycle, except ENDSA 2000.
	  - From each selected PSU, dwellings were sampled systematically.
		Urban: Localities with a population of ≥2,000.
		
		Rural: Localities with a population of <2,000.
			INEI 2012e, 2015a.

	- From each selected dwelling, all children under 5 years of age and women of reproductive age were contacted. - **Data collection**: ENDSA collected information through household and individual questionnaires, anthropometric and blood pressure measurement, and blood specimen.		
Demographic and Health Survey (ENDSA) ENDSA 2000 Jul – Nov 2000. ENDSA 2003-2008: conducted continuously.	- **PSU**: Census township in the last census. - **Stratification**: A stratum was defined by a combinations of department, and departmental capital-other urban-rural area. Within each department, number of conglomerates to be sampled were distributed according to the size of population of each stratum. - **Sampling**: ENDSA started to apply the continuous survey starting from 2003. The first cycle of a continuous survey was implemented between 2003 and 2008. In 2007 and 2008, additional samples were added to estimate selected indicators at departmental level. - From each stratum, PSUs were sampled with probability proportional to size (PPS) of population. - From each selected PSU, conglomerates were sampled with PPS. A conglomerate consisted of an average of 100 dwellings. Selected conglomerates were distributed into 5 survey years systematically. - From each selected conglomerate, dwellings were sampled systematically. - From each selected dwelling, all children under 5 years of age and women of reproductive age were contacted. - **Data collection**: Same as the previous method described in ENDSA 2009-2014.	**Urban**: Localities with a population of ≥2,000. **Rural**: Localities with a population of <2,000.	INEI & Macro International 2001a, INEI & ORC Macro 2007, 2009.

CENAN, Centro Nacional de Alimentación y Nutrición; DEVAN, Dirección Ejecutiva de Vigilancia Alimentaria y Nutricional; INEI, Instituto Nacional de Estadística e Informática.

Appendix D Estimated parameters from the models with age and time as covariates

Table D Estimated parameters from the models with age and time as covariates by country and sex

Mexico, men

Variable	μ parameter			σ parameter			ν parameter			τ parameter			Global test for age group*
	Coeff.	95% CI	p-value	Coeff.	95% CI	p-value	Coeff.	95% CI	p-value	Coeff.	95% CI	p-value	
Age group													
20-24	24.9	(24.6, 25.1)	< 0.001	-1.72	(-1.76, -1.68)	< 0.001	-0.87	(-1.10, -0.64)	< 0.001	0.79	(0.67, 0.90)	< 0.001	
25-29	26.3	(26.0, 26.6)	< 0.001	-1.77	(-1.81, -1.73)	< 0.001	-0.44	(-0.72, -0.16)	< 0.01	0.67	(0.53, 0.81)	< 0.001	
30-34	27.2	(27.0, 27.5)	< 0.001	-1.79	(-1.84, -1.74)	< 0.001	-0.38	(-0.69, -0.06)	0.02	0.45	(0.33, 0.56)	< 0.001	
35-39	27.9	(27.6, 28.2)	< 0.001	-1.76	(-1.81, -1.72)	< 0.001	-0.40	(-0.67, -0.13)	< 0.01	0.47	(0.33, 0.60)	< 0.001	
40-44	28.3	(28.0, 28.5)	< 0.001	-1.87	(-1.91, -1.84)	< 0.001	0.01	(-0.23, 0.25)	0.95	0.46	(0.36, 0.56)	< 0.001	
45-49	27.9	(27.6, 28.2)	< 0.001	-1.82	(-1.87, -1.78)	< 0.001	-0.32	(-0.61, -0.03)	0.03	0.56	(0.43, 0.69)	< 0.001	
50-54	28.0	(27.8, 28.3)	< 0.001	-1.82	(-1.87, -1.77)	< 0.001	-0.25	(-0.54, 0.04)	0.09	0.38	(0.25, 0.51)	< 0.001	
55-59	27.5	(27.2, 27.9)	< 0.001	-1.88	(-1.93, -1.82)	< 0.001	-0.22	(-0.62, 0.18)	0.27	0.53	(0.38, 0.67)	< 0.001	
60-64	27.6	(27.2, 27.9)	< 0.001	-1.90	(-1.95, -1.84)	< 0.001	0.56	(0.15, 0.96)	< 0.01	0.51	(0.34, 0.67)	< 0.001	
65-69	26.8	(26.4, 27.2)	< 0.001	-1.84	(-1.91, -1.78)	< 0.001	0.09	(-0.32, 0.51)	0.66	0.43	(0.28, 0.58)	< 0.001	
Age group × time													
20-24	0.01	(-0.07, 0.10)	0.80	0.01	(-0.01, 0.02)	0.32	-0.10	(-0.17, -0.03)	< 0.01	0.04	(0.00, 0.07)	0.03	< 0.01
25-29	0.10	(0.00, 0.19)	0.04	0.01	(-0.01, 0.02)	0.24	-0.08	(-0.18, 0.02)	0.11	0.02	(-0.02, 0.07)	0.31	< 0.01
30-34	0.09	(0.01, 0.16)	0.02	0.01	(0.00, 0.02)	0.18	-0.07	(-0.17, 0.03)	0.17	0.00	(-0.04, 0.03)	0.86	0.01
35-39	0.11	(0.02, 0.21)	0.02	0.02	(0.00, 0.03)	0.02	0.00	(-0.08, 0.08)	0.98	-0.01	(-0.04, 0.03)	0.76	0.03
40-44	0.14	(0.06, 0.22)	< 0.001	0.00	(-0.01, 0.01)	0.95	0.06	(-0.02, 0.14)	0.14	-0.01	(-0.04, 0.02)	0.38	0.01
45-49	0.07	(-0.02, 0.16)	0.13	0.01	(-0.01, 0.02)	0.35	-0.06	(-0.16, 0.04)	0.22	0.03	(0.00, 0.07)	0.08	0.06
50-54	0.07	(-0.02, 0.16)	0.13	0.01	(0.00, 0.03)	0.18	-0.06	(-0.15, 0.03)	0.22	-0.05	(-0.09, 0.00)	0.03	0.03
55-59	-0.02	(-0.15, 0.11)	0.77	-0.01	(-0.03, 0.01)	0.25	0.01	(-0.12, 0.14)	0.89	-0.04	(-0.09, 0.01)	0.14	0.42
60-64	0.05	(-0.07, 0.17)	0.38	-0.02	(-0.04, 0.00)	0.06	0.06	(-0.05, 0.18)	0.28	-0.02	(-0.07, 0.02)	0.33	0.29
65-69	0.08	(-0.04, 0.20)	0.17	-0.01	(-0.03, 0.01)	0.22	-0.03	(-0.17, 0.10)	0.61	0.00	(-0.06, 0.06)	0.99	0.25

Table D Estimated parameters from the models with age and time as covariates (continued)

Mexico, women

Variable	μ parameter			σ parameter			ν parameter			τ parameter			Global test for age group*
	Coeff.	95% CI	p-value	Coeff.	95% CI	p-value	Coeff.	95% CI	p-value	Coeff.	95% CI	p-value	
Age group													
20-24	24.8	(24.5, 25.0)	< 0.001	-1.64	(-1.68, -1.59)	< 0.001	-0.79	(-1.04, -0.53)	< 0.001	0.65	(0.51, 0.80)	< 0.001	
25-29	26.5	(26.2, 26.8)	< 0.001	-1.64	(-1.67, -1.60)	< 0.001	-0.65	(-0.84, -0.46)	< 0.001	0.70	(0.58, 0.81)	< 0.001	
30-34	27.7	(27.4, 28.0)	< 0.001	-1.67	(-1.70, -1.64)	< 0.001	-0.50	(-0.67, -0.34)	< 0.001	0.66	(0.55, 0.76)	< 0.001	
35-39	28.6	(28.3, 28.8)	< 0.001	-1.71	(-1.75, -1.68)	< 0.001	-0.70	(-0.90, -0.50)	< 0.001	0.56	(0.48, 0.64)	< 0.001	
40-44	29.1	(28.9, 29.4)	< 0.001	-1.72	(-1.75, -1.69)	< 0.001	-0.44	(-0.63, -0.26)	< 0.001	0.57	(0.49, 0.66)	< 0.001	
45-49	29.7	(29.4, 30.0)	< 0.001	-1.71	(-1.75, -1.67)	< 0.001	-0.08	(-0.29, 0.13)	0.47	0.53	(0.42, 0.64)	< 0.001	
50-54	29.8	(29.5, 30.2)	< 0.001	-1.69	(-1.73, -1.65)	< 0.001	-0.38	(-0.62, -0.14)	< 0.01	0.54	(0.39, 0.69)	< 0.001	
55-59	29.7	(29.3, 30.1)	< 0.001	-1.74	(-1.78, -1.70)	< 0.001	-0.12	(-0.38, 0.15)	0.38	0.61	(0.48, 0.74)	< 0.001	
60-64	29.3	(28.9, 29.7)	< 0.001	-1.68	(-1.74, -1.63)	< 0.001	-0.06	(-0.34, 0.22)	0.68	0.49	(0.36, 0.63)	< 0.001	
65-69	28.9	(28.4, 29.3)	< 0.001	-1.71	(-1.77, -1.65)	< 0.001	0.43	(0.10, 0.76)	0.01	0.62	(0.44, 0.79)	< 0.001	
Age group × time													
20-24	0.07	(-0.01, 0.15)	0.11	0.02	(0.01, 0.03)	< 0.01	0.05	(-0.02, 0.12)	0.20	0.00	(-0.04, 0.04)	0.86	0.05
25-29	0.00	(-0.08, 0.09)	0.93	0.02	(0.00, 0.03)	< 0.01	-0.03	(-0.09, 0.04)	0.40	-0.01	(-0.04, 0.03)	0.70	0.09
30-34	0.06	(-0.02, 0.14)	0.13	0.01	(0.00, 0.02)	< 0.01	0.01	(-0.05, 0.06)	0.77	0.00	(-0.03, 0.04)	0.90	0.03
35-39	0.03	(-0.04, 0.10)	0.36	0.00	(-0.01, 0.01)	0.38	-0.06	(-0.13, 0.00)	0.05	0.00	(-0.03, 0.03)	0.88	0.08
40-44	0.04	(-0.04, 0.12)	0.33	0.00	(-0.01, 0.01)	0.75	-0.01	(-0.07, 0.05)	0.80	-0.01	(-0.04, 0.02)	0.57	0.76
45-49	0.09	(0.00, 0.19)	0.06	0.00	(-0.02, 0.01)	0.56	0.03	(-0.03, 0.09)	0.31	-0.02	(-0.05, 0.01)	0.23	0.34
50-54	0.08	(-0.03, 0.19)	0.15	0.01	(-0.01, 0.02)	0.28	-0.02	(-0.11, 0.06)	0.57	0.03	(-0.01, 0.07)	0.19	0.24
55-59	0.07	(-0.05, 0.18)	0.27	0.00	(-0.02, 0.01)	0.63	-0.05	(-0.13, 0.04)	0.26	0.05	(0.01, 0.09)	0.01	0.03
60-64	-0.03	(-0.15, 0.09)	0.60	-0.01	(-0.03, 0.01)	0.24	-0.03	(-0.11, 0.06)	0.53	0.02	(-0.02, 0.07)	0.31	0.58
65-69	0.11	(-0.02, 0.25)	0.09	-0.02	(-0.04, 0.00)	0.06	0.02	(-0.09, 0.12)	0.78	0.06	(0.00, 0.12)	0.04	0.04

Table D Estimated parameters from the models with age and time as covariates (continued)

Colombia, men

Variable	μ parameter			σ parameter			ν parameter			τ parameter			Global test for age group*
	Coeff.	95% CI	p-value	Coeff.	95% CI	p-value	Coeff.	95% CI	p-value	Coeff.	95% CI	p-value	
Age group													
20-24	22.6	(22.5, 22.7)	< 0.001	-1.96	(-1.99, -1.93)	< 0.001	-1.28	(-1.46, -1.10)	< 0.001	0.62	(0.55, 0.70)	< 0.001	
25-29	24.0	(23.9, 24.2)	< 0.001	-1.88	(-1.90, -1.85)	< 0.001	-0.92	(-1.13, -0.72)	< 0.001	0.74	(0.66, 0.82)	< 0.001	
30-34	25.0	(24.8, 25.2)	< 0.001	-1.88	(-1.91, -1.85)	< 0.001	-0.51	(-0.71, -0.31)	< 0.001	0.64	(0.56, 0.72)	< 0.001	
35-39	25.4	(25.3, 25.6)	< 0.001	-1.90	(-1.93, -1.87)	< 0.001	-0.44	(-0.66, -0.21)	< 0.001	0.69	(0.61, 0.78)	< 0.001	
40-44	25.7	(25.5, 25.8)	< 0.001	-1.89	(-1.92, -1.86)	< 0.001	-0.18	(-0.39, 0.03)	0.10	0.59	(0.50, 0.67)	< 0.001	
45-49	25.7	(25.5, 25.8)	< 0.001	-1.89	(-1.92, -1.86)	< 0.001	-0.20	(-0.42, 0.03)	0.09	0.63	(0.54, 0.71)	< 0.001	
50-54	25.6	(25.4, 25.8)	< 0.001	-1.84	(-1.88, -1.81)	< 0.001	-0.15	(-0.39, 0.09)	0.22	0.64	(0.53, 0.74)	< 0.001	
55-59	25.7	(25.5, 25.9)	< 0.001	-1.85	(-1.89, -1.82)	< 0.001	-0.05	(-0.32, 0.22)	0.72	0.60	(0.50, 0.71)	< 0.001	
60-64	25.5	(25.3, 25.7)	< 0.001	-1.87	(-1.91, -1.83)	< 0.001	0.10	(-0.17, 0.37)	0.48	0.67	(0.54, 0.80)	< 0.001	
Age group × time													
20-24	0.08	(0.04, 0.11)	< 0.001	0.01	(0.00, 0.02)	0.11	-0.01	(-0.07, 0.05)	0.79	0.02	(-0.01, 0.04)	0.15	< 0.001
25-29	0.10	(0.06, 0.14)	< 0.001	0.01	(0.00, 0.02)	0.03	-0.03	(-0.09, 0.03)	0.30	0.02	(0.00, 0.04)	0.06	< 0.001
30-34	0.13	(0.08, 0.18)	< 0.001	0.01	(0.00, 0.01)	0.10	0.01	(-0.05, 0.07)	0.77	-0.02	(-0.04, 0.01)	0.13	< 0.001
35-39	0.13	(0.08, 0.18)	< 0.001	0.00	(-0.01, 0.00)	0.34	0.02	(-0.04, 0.08)	0.52	-0.01	(-0.03, 0.02)	0.64	< 0.001
40-44	0.13	(0.08, 0.18)	< 0.001	0.00	(-0.01, 0.01)	0.61	0.05	(-0.01, 0.12)	0.12	-0.05	(-0.08, -0.03)	< 0.001	< 0.001
45-49	0.08	(0.03, 0.14)	< 0.01	0.00	(-0.01, 0.00)	0.36	-0.04	(-0.10, 0.03)	0.24	-0.03	(-0.06, -0.01)	0.01	< 0.001
50-54	0.05	(-0.01, 0.12)	0.09	0.00	(-0.01, 0.01)	0.93	-0.04	(-0.12, 0.04)	0.29	0.00	(-0.04, 0.03)	0.91	0.14
55-59	0.12	(0.06, 0.19)	< 0.001	0.00	(-0.02, 0.01)	0.46	0.01	(-0.08, 0.10)	0.85	-0.01	(-0.05, 0.02)	0.49	< 0.01
60-64	0.08	(0.01, 0.15)	0.02	-0.01	(-0.02, 0.00)	0.17	0.03	(-0.05, 0.11)	0.44	0.01	(-0.03, 0.04)	0.70	0.08

Table D Estimated parameters from the models with age and time as covariates (continued)

Colombia, women

Variable	μ parameter			σ parameter			ν parameter			τ parameter			Global test for age group*
	Coeff.	95% CI	p-value	Coeff.	95% CI	p-value	Coeff.	95% CI	p-value	Coeff.	95% CI	p-value	
Age group													
20-24	22.9	(22.8, 23.0)	< 0.001	-1.82	(-1.85, -1.80)	< 0.001	-1.00	(-1.13, -0.86)	< 0.001	0.60	(0.53, 0.66)	< 0.001	
25-29	24.3	(24.1, 24.4)	< 0.001	-1.78	(-1.80, -1.75)	< 0.001	-0.72	(-0.89, -0.55)	< 0.001	0.62	(0.54, 0.69)	< 0.001	
30-34	25.4	(25.2, 25.5)	< 0.001	-1.79	(-1.81, -1.76)	< 0.001	-0.66	(-0.81, -0.52)	< 0.001	0.58	(0.52, 0.65)	< 0.001	
35-39	26.0	(25.8, 26.2)	< 0.001	-1.77	(-1.79, -1.74)	< 0.001	-0.49	(-0.64, -0.33)	< 0.001	0.59	(0.52, 0.66)	< 0.001	
40-44	26.6	(26.5, 26.8)	< 0.001	-1.75	(-1.77, -1.72)	< 0.001	-0.45	(-0.61, -0.30)	< 0.001	0.52	(0.45, 0.59)	< 0.001	
45-49	27.2	(27.0, 27.3)	< 0.001	-1.77	(-1.79, -1.75)	< 0.001	-0.20	(-0.35, -0.05)	< 0.01	0.57	(0.51, 0.64)	< 0.001	
50-54	27.6	(27.4, 27.8)	< 0.001	-1.74	(-1.77, -1.70)	< 0.001	-0.13	(-0.33, 0.08)	0.22	0.47	(0.39, 0.54)	< 0.001	
55-59	27.9	(27.7, 28.1)	< 0.001	-1.74	(-1.77, -1.71)	< 0.001	0.04	(-0.15, 0.24)	0.67	0.58	(0.50, 0.67)	< 0.001	
60-64	27.6	(27.3, 27.8)	< 0.001	-1.69	(-1.73, -1.65)	< 0.001	-0.10	(-0.36, 0.16)	0.46	0.49	(0.39, 0.60)	< 0.001	
Age group × time													
20-24	0.07	(0.04, 0.11)	< 0.001	0.01	(0.00, 0.01)	0.01	-0.02	(-0.07, 0.02)	0.30	0.00	(-0.02, 0.02)	0.88	< 0.001
25-29	0.12	(0.08, 0.15)	< 0.001	0.01	(0.00, 0.02)	< 0.01	0.03	(-0.01, 0.08)	0.15	0.00	(-0.02, 0.02)	0.75	< 0.001
30-34	0.13	(0.08, 0.17)	< 0.001	0.00	(-0.01, 0.01)	0.61	0.02	(-0.02, 0.06)	0.35	-0.01	(-0.03, 0.00)	0.13	< 0.001
35-39	0.11	(0.07, 0.15)	< 0.001	0.00	(0.00, 0.01)	0.31	0.01	(-0.03, 0.05)	0.66	-0.01	(-0.03, 0.01)	0.27	< 0.001
40-44	0.13	(0.08, 0.18)	< 0.001	0.00	(-0.01, 0.01)	0.95	0.01	(-0.04, 0.05)	0.79	-0.03	(-0.05, -0.01)	< 0.01	< 0.001
45-49	0.05	(0.01, 0.10)	0.03	0.00	(-0.01, 0.00)	0.39	-0.03	(-0.07, 0.02)	0.22	0.01	(-0.01, 0.03)	0.46	0.04
50-54	0.06	(0.00, 0.11)	0.05	0.00	(-0.01, 0.01)	0.52	-0.03	(-0.08, 0.03)	0.32	-0.02	(-0.04, 0.01)	0.18	0.07
55-59	0.10	(0.04, 0.17)	< 0.01	-0.01	(-0.02, 0.00)	0.03	0.02	(-0.04, 0.08)	0.43	0.01	(-0.02, 0.03)	0.65	< 0.01
60-64	0.07	(-0.01, 0.14)	0.08	-0.01	(-0.02, 0.01)	0.25	-0.06	(-0.13, 0.01)	0.10	0.01	(-0.02, 0.04)	0.46	0.02

Table D Estimated parameters from the models with age and time as covariates (continued)

Peru, men

Variable	μ parameter			σ parameter			ν parameter			τ parameter			Global test for age group*
	Coeff.	95% CI	p-value	Coeff.	95% CI	p-value	Coeff.	95% CI	p-value	Coeff.	95% CI	p-value	
Age group													
20-24	23.3	(23.1, 23.5)	< 0.001	-2.02	(-2.07, -1.96)	< 0.001	-0.68	(-1.08, -0.27)	< 0.01	0.46	(0.32, 0.60)	< 0.001	
25-29	24.5	(24.3, 24.7)	< 0.001	-2.02	(-2.07, -1.97)	< 0.001	-1.03	(-1.41, -0.66)	< 0.001	0.53	(0.41, 0.65)	< 0.001	
30-34	25.1	(24.9, 25.3)	< 0.001	-2.00	(-2.05, -1.95)	< 0.001	-0.71	(-1.08, -0.33)	< 0.001	0.54	(0.38, 0.69)	< 0.001	
35-39	25.4	(25.2, 25.7)	< 0.001	-1.92	(-1.98, -1.85)	< 0.001	-0.68	(-1.15, -0.21)	< 0.01	0.45	(0.27, 0.63)	< 0.001	
40-44	25.9	(25.6, 26.1)	< 0.001	-1.97	(-2.02, -1.92)	< 0.001	-0.36	(-0.76, 0.05)	0.08	0.58	(0.44, 0.73)	< 0.001	
45-49	26.0	(25.7, 26.3)	< 0.001	-1.89	(-1.94, -1.85)	< 0.001	-0.39	(-0.70, -0.09)	0.01	0.76	(0.62, 0.90)	< 0.001	
50-54	26.0	(25.7, 26.3)	< 0.001	-1.85	(-1.91, -1.80)	< 0.001	-0.45	(-0.88, -0.03)	0.04	0.67	(0.51, 0.83)	< 0.001	
55-59	25.7	(25.4, 26.1)	< 0.001	-1.91	(-1.96, -1.86)	< 0.001	-0.15	(-0.56, 0.26)	0.48	0.82	(0.63, 1.01)	< 0.001	
60-64	25.6	(25.3, 26.0)	< 0.001	-1.83	(-1.90, -1.76)	< 0.001	-0.41	(-0.90, 0.08)	0.10	0.52	(0.34, 0.70)	< 0.001	
65-69	25.2	(24.9, 25.6)	< 0.001	-1.87	(-1.94, -1.80)	< 0.001	-0.46	(-1.05, 0.13)	0.13	0.57	(0.35, 0.78)	< 0.001	
Age group × time													
20-24	0.18	(0.09, 0.27)	< 0.001	0.02	(-0.01, 0.04)	0.15	0.09	(-0.09, 0.26)	0.34	0.02	(-0.04, 0.09)	0.46	< 0.01
25-29	0.22	(0.12, 0.32)	< 0.001	-0.01	(-0.03, 0.01)	0.48	-0.07	(-0.23, 0.10)	0.43	0.03	(-0.03, 0.08)	0.32	< 0.001
30-34	0.25	(0.15, 0.34)	< 0.001	0.01	(-0.01, 0.03)	0.50	0.10	(-0.07, 0.26)	0.25	-0.02	(-0.09, 0.04)	0.47	< 0.001
35-39	0.30	(0.18, 0.43)	< 0.001	0.03	(0.00, 0.05)	0.07	0.13	(-0.08, 0.34)	0.23	0.02	(-0.06, 0.10)	0.56	< 0.001
40-44	0.25	(0.15, 0.35)	< 0.001	0.01	(-0.01, 0.03)	0.35	-0.01	(-0.19, 0.16)	0.90	0.00	(-0.06, 0.06)	0.99	< 0.001
45-49	0.21	(0.09, 0.33)	< 0.001	-0.01	(-0.03, 0.01)	0.29	0.17	(0.03, 0.31)	0.02	-0.01	(-0.07, 0.05)	0.81	< 0.01
50-54	0.21	(0.09, 0.33)	< 0.001	0.01	(-0.02, 0.03)	0.62	0.00	(-0.19, 0.19)	1.00	-0.07	(-0.14, 0.00)	0.05	< 0.01
55-59	0.13	(-0.02, 0.28)	0.09	-0.01	(-0.03, 0.01)	0.30	-0.04	(-0.22, 0.13)	0.63	0.05	(-0.03, 0.13)	0.22	0.08
60-64	0.37	(0.23, 0.52)	< 0.001	-0.03	(-0.06, 0.00)	0.06	-0.08	(-0.28, 0.12)	0.45	0.04	(-0.04, 0.13)	0.29	< 0.001
65-69	0.34	(0.19, 0.49)	< 0.001	-0.02	(-0.05, 0.01)	0.23	0.27	(0.01, 0.53)	0.04	0.01	(-0.09, 0.10)	0.92	< 0.001

Table D Estimated parameters from the models with age and time as covariates (continued)

Peru, women

Variable	μ parameter			σ parameter			ν parameter			τ parameter			Global test for age group*
	Coeff.	95% CI	p-value	Coeff.	95% CI	p-value	Coeff.	95% CI	p-value	Coeff.	95% CI	p-value	
Age group													
20-24	23.3	(23.1, 23.5)	< 0.001	-1.94	(-1.99, -1.89)	< 0.001	-1.06	(-1.37, -0.74)	< 0.001	0.54	(0.42, 0.66)	< 0.001	
25-29	24.4	(24.2, 24.7)	< 0.001	-1.93	(-1.97, -1.89)	< 0.001	-1.13	(-1.46, -0.81)	< 0.001	0.67	(0.54, 0.80)	< 0.001	
30-34	25.8	(25.6, 26.1)	< 0.001	-1.85	(-1.90, -1.81)	< 0.001	-0.36	(-0.67, -0.04)	0.03	0.62	(0.49, 0.75)	< 0.001	
35-39	26.5	(26.3, 26.8)	< 0.001	-1.84	(-1.88, -1.80)	< 0.001	-0.50	(-0.79, -0.22)	< 0.001	0.60	(0.50, 0.71)	< 0.001	
40-44	26.7	(26.5, 26.9)	< 0.001	-1.88	(-1.92, -1.84)	< 0.001	-0.68	(-0.97, -0.40)	< 0.001	0.55	(0.44, 0.66)	< 0.001	
45-49	27.0	(26.8, 27.3)	< 0.001	-1.81	(-1.85, -1.77)	< 0.001	-0.30	(-0.53, -0.07)	0.01	0.62	(0.51, 0.72)	< 0.001	
50-54	27.2	(26.9, 27.5)	< 0.001	-1.72	(-1.76, -1.67)	< 0.001	-0.02	(-0.33, 0.29)	0.91	0.55	(0.42, 0.68)	< 0.001	
55-59	27.0	(26.6, 27.3)	< 0.001	-1.72	(-1.78, -1.67)	< 0.001	-0.13	(-0.42, 0.15)	0.36	0.49	(0.35, 0.64)	< 0.001	
60-64	26.5	(26.1, 26.9)	< 0.001	-1.70	(-1.75, -1.66)	< 0.001	0.11	(-0.20, 0.42)	0.48	0.90	(0.73, 1.07)	< 0.001	
65-69	26.0	(25.5, 26.5)	< 0.001	-1.66	(-1.72, -1.60)	< 0.001	-0.15	(-0.46, 0.16)	0.34	0.69	(0.50, 0.89)	< 0.001	
Age group × time													
20-24	0.09	(0.01, 0.17)	0.04	0.02	(0.00, 0.04)	0.10	0.06	(-0.09, 0.20)	0.43	-0.03	(-0.08, 0.02)	0.30	0.12
25-29	0.06	(-0.04, 0.17)	0.22	0.02	(0.00, 0.04)	0.04	-0.18	(-0.33, -0.03)	0.02	0.05	(-0.01, 0.11)	0.08	< 0.001
30-34	0.20	(0.10, 0.31)	< 0.001	0.01	(-0.01, 0.03)	0.42	0.03	(-0.10, 0.17)	0.63	0.00	(-0.06, 0.05)	0.90	< 0.01
35-39	0.23	(0.12, 0.34)	< 0.001	-0.01	(-0.03, 0.01)	0.35	0.03	(-0.09, 0.16)	0.58	-0.05	(-0.10, 0.00)	0.04	< 0.001
40-44	0.25	(0.15, 0.35)	< 0.001	0.01	(0.00, 0.03)	0.15	0.00	(-0.12, 0.12)	0.95	-0.02	(-0.06, 0.03)	0.46	< 0.001
45-49	0.21	(0.09, 0.33)	< 0.001	-0.01	(-0.02, 0.01)	0.37	0.01	(-0.09, 0.11)	0.90	0.00	(-0.04, 0.05)	0.85	< 0.01
50-54	0.07	(-0.06, 0.21)	0.29	0.00	(-0.02, 0.02)	0.83	0.05	(-0.09, 0.18)	0.48	-0.05	(-0.11, 0.01)	0.08	0.32
55-59	0.32	(0.18, 0.47)	< 0.001	-0.02	(-0.04, 0.00)	0.12	0.08	(-0.04, 0.20)	0.21	0.00	(-0.06, 0.06)	0.97	< 0.001
60-64	0.33	(0.15, 0.51)	< 0.001	0.00	(-0.02, 0.02)	0.82	-0.02	(-0.16, 0.12)	0.78	-0.07	(-0.15, 0.00)	0.05	< 0.001
65-69	0.37	(0.17, 0.57)	< 0.001	0.01	(-0.01, 0.04)	0.32	0.02	(-0.11, 0.16)	0.76	0.01	(-0.08, 0.09)	0.86	< 0.01

* *P*-values are from 4 *df* likelihood ratio tests. The null hypothesis is: $\beta_{2j}^{\mu} = 0, \beta_{2j}^{\sigma} = 0, \beta_{2j}^{\nu} = 0, \beta_{2j}^{\tau} = 0$ where β_{2j}^{i} is the coefficient for the interaction term between age group j and time for the parameter $i = (\mu, \sigma, \nu, \tau)$.

Appendix E Methods used for wealth index construction

The overall procedure was as follows. Household wealth indices were constructed from variables on dwelling characteristics, household services and assets for each survey following an established method by the DHS Program (Rutstein and Kiersten 2004, Rutstein 2008, IFC International 2015, Fry et al. 2014). Household wealth score was calculated with principal component analysis (PCA). The first principal component obtained from the included variables was taken as the underlying wealth score, and household wealth score was calculated with the PCA factor weights, separately for urban and rural households. Urban and rural wealth scores were combined into a national score according to the method described by the DHS program (Rutstein 2008).

First, variables to be included in PCA were selected referring to the DHS guideline (Rutstein 2008) and the variables used to construct wealth index for the DHS conducted in Colombia, Peru, and other countries that conducted surveys recently (The DHS Program n.d.). The included variables were number of people per bedroom, use of domestic servants, house type and ownership, construction materials (wall, floor, and roof), available services (water, electricity, sewage system, cooking fuels), household items, equipment and vehicles, (and number of animal stock and land area for Peruvian survey data). Missing values for binary variables were replaced with zero; hence, one (1) indicates if the household has a “yes” answer to the question and zero (0) indicates if the households has a “no” answers, don’t know, not applicable response, or missing. Missing values for non-binary variables (i.e., discrete continuous variables) were replaced with their median values. However, such replacements were very small (< 0.16%, 0.00%, < 0.01% at maximum per variable in Mexican, Colombian, and Peruvian data, respectively; data not shown). Households that did not have any data about dwelling characteristics (type and construction materials) or household items were not included in wealth index construction, and wealth score were not calculated for such households. The complete list of variables that were used in constructing wealth index were shown in Table E-1.

Second, the first principal component was taken as the underlying wealth score, and household wealth scores were calculated with the obtained PCA factor weights, separately for urban, rural, and all households. Animal livestock and land area, which are thought to have different relationships with the underlying economic status in urban and rural areas, were not used to construct the common wealth score that was calculated using all households. For example, having chickens in urban areas may indicate a poor status whereas it can indicate wealthy status in rural areas (Rutstein 2008). The results of PCA from the Colombian ENDS/ENSIS 2010 data are presented in Table E-2 as an example.

Third, urban and rural wealth scores were combined to produce a national score. The urban and rural wealth scores were regressed on the common wealth score separately with linear regression. Then, the estimated coefficients were used to map the urban and rural wealth scores on the national score. This 2-stage procedure is implemented, instead of constructing one national wealth score at one time, in order to distinguish the extremely poor from the poor (Rutstein 2008). Urban and rural wealth scores were calculated separately using variables that could distinguish population's wealth status in each area; and urban and rural scores were mapped onto a single national wealth score to make these 2 scores comparable.

Lastly, quartiles of wealth scores were calculated from the national wealth score. The quartiles were of population wealth scores, not of household wealth scores. Hence, the quartile cutoffs were calculated using weighted cumulative percentages of the national wealth score, where the weights were obtained by the product of household weights and number of household members. Quartiles were used, instead of commonly used quintiles, to reduce the number of variables in the model in comparison to the variable for education that have 3 categories. Means of variables by wealth quartile were calculated to make sure that the procedure was correctly implemented. An example from the Colombian ENDS/ENSIN 2010 data is presented in Table E-3.

Table E-1 List of variables included in PCA

Mexico, ENSANUT 2006

- Number of people per bedroom
- Domestic servant in the household
- Type of house
  - Own
  - Rented (alquilada)
  - Lent (prestada)
  - Given/donated
  - Wall material
  - Cement, block, stone, wood
  - Adobe
  - Plant, mud, natural mat
  - Cardboard, plastic sheet
  - Lamina
  - Other material
- Roofing material
  - Cardboard, rubber, fabric, tires
  - Cardboard sheets (Lámina de cartón)
  - Palm, shingles or wood
  - Metallic, fiberglass, or plastic sheet
  - Asbestos sheet
  - Reed, bamboo
  - Roof tiles
  - Concrete blocks
  - Brick, breeze block
  - Block
  - Floor material
  - Earth
  - Cement
  - Tiles, wood, other covering material
- Main source of water
  - Piped water, inside
  - Piped water, outside
  - Piped water, public
  - Deep well
  - Protected shallow well
  - Unprotected shallow well
  - Rain water
  - Pond, river water
  - Coco water
- Water for flushing toilet
- Toilet type
  - Toilet with drainage
  - Toilet with septic tank
  - Latrine with drainage
  - Pit latrine with cover
  - Pit latrine without cover
  - Latrine with bucket
  - Open air
  - Drainage to nature
  - Other (Pit)
  - Other (Latrine)
  - Other
  - No toilet, or use neighbors'
- Exclusive use of toilet
- Drainage system
  - Connected to the street
  - Connected to a septic tank
  - Connected to a river, lake, ditch
  - No drainage
- Cooking fuel
  - Gas
  - Electricity
  - Kerosene
  - Mineral coal
  - Charcoal
  - Firewood
  - Agricultural residue
  - Bush, grass
  - Other
  - Do not cook
- Cooking place
  - Kitchen room
  - Kitchen room under construction
  - Other room
  - Open air
- Electricity for lighting
- Heating system
- Air conditioning
- Household items
  - Other real estate
  - Car
  - Truck
  - Motorcycle, bicycle
  - Boat
  - Black & white TV
  - Color TV
  - Parabolic antenna
  - Radio
  - Stereo
  - Refrigerator
  - Gas stove
  - Other type of stove
  - Washing machine
  - Boiler
  - Computer
  - Microwave
  - Telephone
  - Mixer
  - VCR
  - Fan
  - Tractor

Table E-1 List of variables included in PCA (continued)

Mexico, ENSANUT 2012

Number of people per bedroom
Domestic servant in the household
Type of house
- Own with owner in the household
- Own without owner in the household
- Own, owner not known
- Rent
- Other

Wall material
- Waste material
- Cardboard sheets (Lámina de cartón)
- Asbestos or metal sheets
- Reed, bamboo, palm
- Mud or straw
- Wood
- Adobe
- Brick, block, stone, cement, concrete

Roof material
- Waste material
- Cardboard sheets (Lámina de cartón)
- Metallic sheets
- Asbestos sheets
- Palm or straw
- Wood or shingles
- "Terrado con vigueria"
- Roof tiles
- Concrete slab or joist with vaults

Floor material
- Earth
- Cement
- Wooden tiles or other covering material

Main source of water
- Piped water, inside
- Piped water, outside
- Piped water, public
- Piped water, from neighbor
- Coco water
- Well, river, lake, stream, other

Toilet type
- Flushing toilet, private
- Non-flushing toilet, private
- Latrine, private
- Flushing toilet, shared
- Non-flushing toilet, shared
- Latrine, shared
- No toilet

Drainage system
- Public drainpipe system
- Septic tank
- Connected to ditch
- Connected to river, lake or sea
- No drainage

Cooking fuel
- Gas cylinder or tank
- Natural gas or gas pipe
- Firewood
- Charcoal
- Electricity
- Other

Cooking place
- Kitchen room
- Kitchen place under construction
- Other room
- Open air

Electricity for lighting
Heating system
Household item
- Other house, real estate, land
- Car
- Truck
- Motorcycle, scooter
- Boat, canoe
- TV
- Cable TV
- Radio
- Stereo
- Iron
- Mixer
- Refrigerator
- Gas stove
- Other stove
- Washing machine or dryer
- Boiler
- Computer
- Internet
- Microwave
- Landline telephone
- Mobile phone
- Air conditioner

Table E-1 List of variables included in PCA (continued)

Colombia, ENDS/ENSIN 2005

No. of people per bedroom
Domestic servant in household
Type of house
- House
- Apartment
- Room in house for rent
- Room in other type of structure
- Other

Floor material
- Earth, sand
- Wood planks
- Cement, gravel
- Marmol, parquet, polished wood
- Rug, carpet
- Ceramic tiles, vinyl, bricks
- Other

Wall material
- Bamboo, straw, other plants
- Adobe
- Bamboo with mud plaster
- Planks
- Compact dirt or mud
- Zinc, canvas, plastics
- Bricks, polished wood, premanufactured material
- No walls
- Other

Source of non-drinking water
- Piped water from utility company
- Piped water from rural system
- Public tap
- Open well with sump pump
- Open well without sump pump
- River, stream, spring
- Rain water
- Tanker truck
- Water in drums, big cans
- Bottled water
- Other

Toilet typw
- Toilet connected to sewer, private
- Toilet connected to septic well, private
- Toilet connected to plot/yard, private
- Traditional pit toilet, private
- Traditional toilet to sea/river, private
- Other, private
- Toilet connected to sewer, shared
- Toilet connected to septic well, shared
- Toilet connected to plot/yard, shared
- Traditional pit toilet, shared
- Traditional toilet to sea/river, shared
- Other, shared
- No toilet facility

Number of toilets
Place of toilet
- Outside dwelling area
- Within plot but outside dwelling
- Inside dwelling
- Other

Cooking fuel
- Natural gas
- Propane gas
- Kerosene/oil/cocinol/diesel/gasoline/alcohol
- Electricity
- Firewood, charcoal
- Mineral coal
- Disposable material
- Do not cook

Has electricity
Access to natural gas
Access to piped water
Access to sewer
Access to garbage collection
Landline telephone, exclusive use
Landline telephone, shared
Household item
- Radio
- TV
- Refrigerator
- Motorcycle, scooter
- Car, truck
- Shower
- Blender
- Stereo
- Washing machine
- DVD player
- Computer
- Internet
- Electric or gas stove
- Electric or gas oven
- Microwave oven
- Vaccum cleaner, polisher
- Electric or gas water heater
- Air conditioning, heater unit
- VHS, betamax
- Fan

Table E-1 List of variables included in PCA (continued)

Colombia, ENDS/ENSIN 2010

- Number of people per bedroom
- Domestic servant in the household
- Type of house
  - House, owned
  - House, rented
  - Apartment, owned
  - Apartment, rented
  - Room in house for rent
  - Room in other type of structure
  - Indigenous dwelling
  - Other
- Floor material
  - Earth/sand
  - Wood planks
  - Cement/gravel
  - Rug, carpet
  - Polished wood, parquet
  - Marble
  - Ceramic tiles, vinyl, bricks
  - Other
- Wall material
  - Bamboo, straw, other plants
  - Bamboo without plaster
  - Adobe
  - Bamboo with mud plaster
  - Planks
  - Zinc, canvas, plastics
  - Bricks/polished wood/pre-manufactured material
  - Prefabricated material
  - Other
  - No walls
- Source of non-drinking water
  - Piped water from utility company
  - Piped water from rural system
  - Public tap
  - Open well with sump pump
  - Open well without sump pump
  - River/stream/spring
  - Rain water
  - Tanker truck
  - Water in drums/big cans
  - Bottled water
  - Other
- Toilet type
  - Toilet connected to sewer, private
  - Toilet connected to septic well, private
  - Toilet connected to plot/yard, private
  - Traditional pit toilet, private
  - Traditional toilet to sea/river, private
  - Other, unshared
  - Toilet connected to sewer, shared
  - Toilet connected to septic well, shared
  - Toilet connected to plot/yard, shared
  - Traditional pit toilet, shared
  - Traditional toilet to sea/river, shared
  - Other, shared
  - No toilet facility
- Number of toilets
- Place of toilet
  - Outside dwelling area
  - Within plot but outside dwelling
  - Inside dwelling
  - Other
- Cooking fuel
  - Natural gas
  - Propane gas
  - Kerosene/oil/cocinol/diesel/gasoline/alcohol
  - Electricity
  - Firewood, charcoal
  - Mineral coal
  - Do not cook
  - Other
- Has electricity
- Access to natural gas
- Access to piped water
- Access to sewer
- Access to garbage collection
- Landline, exclusive use
- Landline, shared with others
- Household items
  - Mobile telephone
  - Radio
  - Television
  - Refrigerator
  - Bicycle
  - Motorcycle/scooter
  - Car/truck
  - Shower
  - Canoe

Table E-1 List of variables included in PCA (continued)

Peru, ENAHO-MONIN 2007-8

- No. people per bedroom
- Domestic servant in household
- Type of house
  - Independent house
  - Apartment
  - Villa
  - Tenement
  - Hut or cabin
  - Makeshift house
  - Place not meant for habitation
  - Other
- Wall material
  - Brick or cement block
  - Stone with cement or lime
  - Adobe
  - Rammed earth
  - Quincha (cane with mud)
  - Stones with mud
  - Wood
  - Matting
  - Other material
- Floor material
  - Parquet, polished wood
  - Asphalt or vinyl sheets
  - Tiles, terrazzo
  - Wood decking
  - Cement
  - Dirt
  - Other material
- Roofing material
- Concrete
  - Wood
  - Roof tiles
  - Metal or fiber cement sheets
  - Cane or matting with mud
  - Matting
  - Straw, palm leaves
  - Other material
- House ownership
  - Own with title
  - Own without title or don't know
  - Rent
  - Handed over
  - Other form
- Water supply
  - Piped water inside the house
  - Piped water outside the house
  - Water basin for public use
  - Water truck
  - Well
  - River, canal, spring
  - Other type 1
  - Other type 2
- Toilet type
  - Toilet inside the house
  - Toilet outside the house
  - Septic tank
  - Latrine, cesspit
  - River, canal
  - Do not have
- Lighting
  - Electricity with meter
  - Electricity without meter
  - Kerosene
  - Oil or gas
  - Candle
  - Generator
  - Other
  - Do not use
- Cooking fuel
  - Electricity
  - LPG gas
  - Natural gas
  - Kerosene
  - Charcoal
  - Firewood
  - Other
  - Do not cook
- Household item
  - Landline telephone
  - Mobile phone
  - Cable TV
  - Internet
  - Radio
  - Color TV
  - Black & white TV
  - Stereo
  - DVD
  - Video recorder
  - Computer
  - Iron
  - Mixer
  - Gas stove
  - Kerosene stove
  - Refrigerator
  - Washing machine
  - Microwave
  - Sewing machine
  - Bicycle
  - Car
  - Motorcycle
  - Tricycle
  - Motorcycle taxi
  - Truck
- Animal livestock
  - Livestock 3100
  - Livestock 3200
  - Livestock 3300
  - Livestock 3400
  - Livestock 3500
  - Livestock 3600
  - Livestock 3700
  - Livestock 3801
  - Livestock 3802
  - Livestock 3900
- Total agricultural land area (ha)

Table E-1 List of variables included in PCA (continued)

Peru, ENAHO-MONIN 2012-3

- No. people per bedroom
- Domestic servant in household
- Type of house
    - Independent house
    - Apartment
    - Villa
    - Tenement
    - Hut or cabin
    - Makeshift house
    - Place not meant for habitation
- Wall material
    - Brick or cement block
    - Stone with cement or lime
    - Adobe
    - Rammed earth
    - Quincha (cane with mud)
    - Stones with mud
    - Wood
    - Matting
    - Other material
    - Floor material
    - Parquet, polished wood
    - Asphalt or vinyl sheets
    - Tiles, terrazzo
    - Wood decking
    - Cement
    - Dirt
    - Other material
- Roofing material
    - Concrete
    - Wood
    - Roof tiles
    - Metal or fiber cement sheets
    - Cane or matting with mud
    - Matting
    - Straw, palm leaves
    - Other material
- House ownership
    - Own with title
    - Own without title or don't know
    - Rent
    - Handed over
    - Other form
- Water supply
    - Piped water inside the house
    - Piped water outside the house
    - Water basin for public use
    - Water truck
    - Well
    - River, canal, spring
    - Other type 1
    - Other type 2
- Toilet type
    - Toilet inside the house
    - Toilet outside the house
    - Latrine
    - Septic tank
    - Cesspit
    - River, canal
    - Other
    - Do not have
- Lighting
    - Electricity with meter
    - Electricity without meter
    - Kerosene
    - Oil or gas
    - Candle
    - Generator
    - Other
    - Do not use
- Cooking fuel
    - Electricity
    - LPG gas
    - Natural gas
    - Kerosene
    - Charcoal
    - Firewood
    - Other
    - Do not cook
- Household item
    - Landline telephone
    - Mobile phone
    - Cable TV
    - Internet
    - Radio
    - Color TV
    - Black & white TV
    - Stereo
    - DVD
    - Video recorder
    - Computer
    - Iron
    - Mixer
    - Gas stove
    - Kerosene stove
    - Refrigerator
    - Washing machine
    - Microwave
    - Sewing machine
    - Bicycle
    - Car
    - Motorcycle
    - Tricycle
    - Motorcycle taxi
    - Truck
- Animal livestock
    - Livestock 3100
    - Livestock 3200
    - Livestock 3300
    - Livestock 3400
    - Livestock 3500
    - Livestock 3600
    - Livestock 3700
    - Livestock 3801
    - Livestock 3802
    - Livestock 3900
    - Livestock 4113
- Total agricultural land area (ha)

Table E-2 PCA factor weights (example from Colombia, ENDS/ENSIN 2010)

Variable	No. miss-ing	Min value	Max value	All households			Urban households			Rural households		
				Mean	SD	Factor weight	Mean	SD	Factor weight	Mean	SD	Factor weight
Number of people per bedroom	0	0	21	1.733	1.345	-0.114	1.614	1.227	-0.156	2.023	1.557	-0.125
Domestic servant in the household	0	0	1	0.008	0.090	0.022	0.010	0.099	0.030	0.004	0.064	0.016
Type of house												
House, owned	0	0	1	0.549	0.498	-0.057	0.497	0.500	-0.007	0.677	0.468	-0.067
House, rented	0	0	1	0.297	0.457	-0.001	0.301	0.459	-0.042	0.287	0.452	0.056
Apartment, owned	0	0	1	0.034	0.181	0.063	0.047	0.212	0.075	0.002	0.042	0.023
Apartment, rented	0	0	1	0.093	0.290	0.069	0.125	0.331	0.050	0.015	0.123	0.043
Room in house for rent	0	0	1	0.012	0.110	-0.003	0.015	0.123	-0.039	0.004	0.066	0.008
Room in other type of structure	0	0	1	0.011	0.105	-0.005	0.013	0.113	-0.034	0.006	0.079	0.007
Indigenous dwelling	0	0	1	0.000	0.018	-0.017				0.001	0.034	-0.028
Other	0	0	1	0.003	0.059	-0.019	0.002	0.043	-0.023	0.007	0.086	-0.005
Floor material												
Earth/sand	0	0	1	0.073	0.260	-0.144	0.035	0.183	-0.161	0.164	0.371	-0.174
Wood planks	0	0	1	0.074	0.262	-0.109	0.040	0.195	-0.088	0.157	0.364	-0.140
Cement/gravel	0	0	1	0.435	0.496	-0.066	0.387	0.487	-0.138	0.553	0.497	0.134
Rug, carpet	0	0	1	0.004	0.060	0.020	0.005	0.070	0.023	0.000	0.020	0.011
Polished wood, parquet	0	0	1	0.007	0.085	0.013	0.009	0.092	0.018	0.004	0.064	-0.001
Marble	0	0	1	0.003	0.051	0.017	0.004	0.061	0.021	0.000	0.008	0.004
Ceramic tiles, vinyl, bricks	0	0	1	0.404	0.491	0.194	0.521	0.500	0.219	0.121	0.327	0.148
Other	0	0	1	0.000	0.013	0.000	0.000	0.013	0.002	0.000	0.014	0.001
Wall material												
Bamboo, straw, other plants	0	0	1	0.011	0.102	-0.044	0.006	0.077	-0.044	0.022	0.145	-0.058
Bamboo without plaster	0	0	1	0.017	0.130	-0.062	0.007	0.081	-0.059	0.043	0.203	-0.056
Adobe	0	0	1	0.020	0.141	-0.045	0.009	0.094	-0.026	0.048	0.213	-0.019
Bamboo with mud plaster	0	0	1	0.035	0.185	-0.054	0.020	0.140	-0.052	0.073	0.260	-0.018
Planks	0	0	1	0.127	0.333	-0.178	0.058	0.233	-0.189	0.293	0.455	-0.190
Zinc, canvas, plastics	0	0	1	0.006	0.075	-0.029	0.005	0.069	-0.056	0.008	0.088	-0.023
Bricks/polished wood/premanufac. mtl.	0	0	1	0.775	0.417	0.215	0.886	0.317	0.206	0.507	0.500	0.233
Prefabricated material	0	0	1	0.008	0.089	0.011	0.009	0.096	0.006	0.004	0.067	0.015
Other	0	0	1	0.000	0.021	-0.009	0.000	0.015	-0.005	0.001	0.032	-0.013
No walls	0	0	1	0.000	0.016	-0.010	0.000	0.010	-0.007	0.001	0.026	-0.014
Source of non-drinking water												
Piped water from utility company	0	0	1	0.686	0.464	0.241	0.894	0.307	0.195	0.181	0.385	0.151
Piped water from rural system	0	0	1	0.122	0.327	-0.105	0.016	0.126	-0.070	0.378	0.485	0.099

Table E-2 PCA factor weights (example from Colombia, ENDS/ENSIN 2010) (continued)

Variable	No. miss-ing	Min value	Max value	All households			Urban households			Rural households		
				Mean	SD	Factor weight	Mean	SD	Factor weight	Mean	SD	Factor weight
Public tap	0	0	1	0.001	0.030	-0.011	0.001	0.028	-0.025	0.001	0.035	-0.006
Open well with sump pump	0	0	1	0.055	0.227	-0.045	0.046	0.210	-0.089	0.075	0.264	0.003
Open well without sump pump	0	0	1	0.026	0.158	-0.080	0.007	0.086	-0.058	0.070	0.255	-0.083
River/stream/spring	0	0	1	0.062	0.242	-0.139	0.003	0.057	-0.046	0.205	0.404	-0.136
Rain water	0	0	1	0.036	0.185	-0.089	0.021	0.142	-0.109	0.072	0.258	-0.108
Tanker truck	0	0	1	0.004	0.061	-0.011	0.003	0.055	-0.031	0.006	0.074	0.012
Water in drums/big cans	0	0	1	0.003	0.050	-0.014	0.003	0.053	-0.037	0.002	0.044	-0.008
Bottled water	0	0	1	0.000	0.016	-0.003	0.000	0.017	-0.007	0.000	0.014	-0.005
Other	0	0	1	0.006	0.079	-0.036	0.005	0.070	-0.068	0.010	0.098	-0.032
Toilet												
Toilet connected to sewer, private	0	0	1	0.585	0.493	0.239	0.754	0.431	0.256	0.175	0.380	0.177
Toilet connected to septic well, private	0	0	1	0.183	0.387	-0.127	0.081	0.273	-0.145	0.432	0.495	0.063
Toilet connected to plot/yard, private	0	0	1	0.035	0.183	-0.059	0.016	0.127	-0.087	0.080	0.271	0.008
Traditional pit toilet, private	0	0	1	0.012	0.109	-0.064	0.003	0.050	-0.047	0.035	0.184	-0.071
Traditional toilet to sea/river, private	0	0	1	0.005	0.073	-0.020	0.005	0.068	-0.050	0.007	0.083	0.005
Other, unshared	0	0	1	0.001	0.022	-0.007	0.000	0.017	-0.009	0.001	0.032	-0.003
Toilet connected to sewer, shared	0	0	1	0.082	0.274	0.012	0.105	0.307	-0.068	0.025	0.157	0.052
Toilet connected to septic well, shared	0	0	1	0.021	0.145	-0.047	0.012	0.109	-0.079	0.044	0.205	0.003
Toilet connected to plot/yard, shared	0	0	1	0.005	0.069	-0.022	0.004	0.060	-0.050	0.007	0.086	0.003
Traditional pit toilet, shared	0	0	1	0.001	0.037	-0.018	0.001	0.026	-0.027	0.003	0.055	-0.012
Traditional toilet to sea/river, shared	0	0	1	0.001	0.027	-0.008	0.001	0.024	-0.020	0.001	0.034	0.004
Other, shared	0	0	1	0.000	0.011	-0.003	0.000	0.009	-0.006	0.000	0.014	0.002
No toilet facility	0	0	1	0.069	0.253	-0.171	0.019	0.137	-0.148	0.189	0.392	-0.246
Number of toilets	0	0	6	1.174	0.661	0.160	1.288	0.666	0.152	0.899	0.561	0.233
Place of toilet												
Outside dwelling area	0	0	1	0.010	0.098	-0.021	0.007	0.085	-0.025	0.015	0.123	-0.010
Within plot but outside dwelling	0	0	1	0.194	0.395	-0.121	0.137	0.344	-0.176	0.332	0.471	-0.008
Inside dwelling	0	0	1	0.727	0.445	0.209	0.836	0.370	0.224	0.463	0.499	0.203
Other	0	0	1	0.000	0.015	-0.001	0.000	0.015	-0.004	0.000	0.016	0.003
Cooking fuel												
Natural gas	0	0	1	0.388	0.487	0.197	0.530	0.499	0.215	0.045	0.208	0.083
Propane gas	0	0	1	0.371	0.483	-0.014	0.360	0.480	-0.121	0.396	0.489	0.202
Kerosene/oil/cocinol/diesel/gasoline/alcohol	0	0	1	0.004	0.061	-0.015	0.004	0.062	-0.038	0.004	0.059	-0.003
Electricity	0	0	1	0.031	0.173	-0.003	0.036	0.186	-0.042	0.020	0.139	0.024

Table E-2 PCA factor weights (example from Colombia, ENDS/ENSIN 2010) (continued)

Variable	No. missing	Min value	Max value	All households			Urban households			Rural households		
				Mean	SD	Factor weight	Mean	SD	Factor weight	Mean	SD	Factor weight
Firewood, charcoal	0	0	1	0.174	0.379	-0.225	0.036	0.186	-0.159	0.507	0.500	-0.241
Mineral coal	0	0	1	0.003	0.050	-0.009	0.001	0.033	-0.016	0.006	0.077	0.016
Do not cook	0	0	1	0.030	0.170	-0.012	0.033	0.178	-0.050	0.022	0.148	0.001
Other	0	0	1	0.000	0.015	-0.006	0.000	0.005	-0.004	0.001	0.027	0.001
Other available services												
Has electricity	0	0	1	0.950	0.218	0.143	0.991	0.094	0.061	0.851	0.356	0.196
Access to natural gas	0	0	1	0.404	0.491	0.197	0.550	0.498	0.212	0.053	0.224	0.081
Access to piped water	0	0	1	0.816	0.387	0.205	0.925	0.264	0.185	0.555	0.497	0.213
Access to sewer	0	0	1	0.680	0.467	0.254	0.869	0.337	0.250	0.220	0.414	0.188
Access to garbage collection	0	0	1	0.770	0.421	0.244	0.970	0.172	0.155	0.286	0.452	0.207
Landline, exclusive use	0	0	1	0.278	0.448	0.168	0.375	0.484	0.190	0.043	0.203	0.105
Landline, shared with others	0	0	1	0.017	0.127	0.020	0.022	0.146	0.002	0.003	0.059	0.018
Household items												
Mobile telephone	0	0	1	0.878	0.327	0.123	0.925	0.263	0.090	0.765	0.424	0.155
Radio	0	0	1	0.718	0.450	0.100	0.751	0.433	0.145	0.639	0.480	0.095
Television	0	0	1	0.873	0.333	0.168	0.941	0.235	0.131	0.709	0.454	0.205
Refrigerator	0	0	1	0.709	0.454	0.187	0.802	0.398	0.182	0.484	0.500	0.226
Bicycle	0	0	1	0.325	0.468	0.055	0.354	0.478	0.041	0.254	0.435	0.077
Motorcycle/scooter	0	0	1	0.223	0.416	0.040	0.240	0.427	0.020	0.180	0.384	0.093
Car/truck	0	0	1	0.102	0.302	0.087	0.125	0.331	0.112	0.045	0.208	0.077
Shower	0	0	1	0.716	0.451	0.218	0.823	0.382	0.243	0.459	0.498	0.241
Canoe	0	0	1	0.039	0.194	-0.110	0.011	0.105	-0.054	0.108	0.310	-0.154
Total number of households				51,447			36,412			15,035		
Eigenvalue, 1st component					9.55			6.9			7.0	
Proportion of variance explained					0.118			0.086			0.086	

* Results after excluding 341 records without any data on household item or household characteristics. Then, missing values were replaced with zero if the variable is binary or median if the variable is non-binary for PCA.

† After this PCA analysis, the common (all household) score was regressed on the urban and rural scores separately in order to obtain formulas to map the rural and urban scores onto the national score. The formulas used to calculate the national score were:

National wealth score for urban households = 1.4004 + 0.7559 × urban wealth score

National wealth score for rural households =−3.3913 + 0.9710 × rural wealth score

Table E-3 Variable means by wealth quartile (example from Colombia, ENDS/ENSIN 2010)

Variable	Mean within wealth quartile				Total mean
	Q1	Q2	Q3	Q4	
Number of people per bedroom	2.2	1.9	1.5	1.0	1.6
Domestic servant in the household	0.00	0.00	0.01	0.02	0.01
Type of house					
House, owned	0.67	0.45	0.45	0.49	0.51
House, rented	0.29	0.37	0.30	0.15	0.28
Apartment, owned	0.00	0.01	0.04	0.18	0.06
Apartment, rented	0.01	0.11	0.19	0.18	0.13
Room in house for rent	0.01	0.03	0.00	0.00	0.01
Room in other type of structure	0.01	0.03	0.01	0.00	0.01
Indigenous dwelling	0.00	0.00	0.00	0.00	0.00
Other	0.01	0.00	0.00	0.00	0.00
Floor material					
Earth/sand	0.23	0.01	0.00	0.00	0.06
Wood planks	0.10	0.04	0.02	0.00	0.04
Cement/gravel	0.60	0.63	0.30	0.01	0.38
Rug, carpet	0.00	0.00	0.00	0.02	0.01
Polished wood, parquet	0.00	0.00	0.01	0.02	0.01
Marble	0.00	0.00	0.00	0.01	0.00
Ceramic tiles, vinyl, bricks	0.07	0.30	0.66	0.93	0.50
Other	0.00	0.00	0.00	0.00	0.00
Wall material					
Bamboo, straw, other plants	0.03	0.01	0.00	0.00	0.01
Bamboo without plaster	0.06	0.01	0.00	0.00	0.02
Adobe	0.07	0.02	0.00	0.00	0.02
Bamboo with mud plaster	0.09	0.03	0.00	0.00	0.03
Planks	0.22	0.02	0.00	0.00	0.06
Zinc, canvas, plastics	0.01	0.00	0.00	0.00	0.00
Bricks/polished wood/premanufac. mtl.	0.51	0.90	0.97	0.99	0.85
Prefabricated material	0.00	0.01	0.02	0.01	0.01
Other	0.00	0.00	0.00	0.00	0.00
No walls	0.00	0.00	0.00	0.00	0.00
Source of non-drinking water					
Piped water from utility company	0.22	0.85	0.99	1.00	0.78
Piped water from rural system	0.36	0.12	0.00	0.00	0.11
Public tap	0.00	0.00	0.00	0.00	0.00
Open well with sump pump	0.05	0.01	0.00	0.00	0.02
Open well without sump pump	0.08	0.00	0.00	0.00	0.02
River/stream/spring	0.21	0.01	0.00	0.00	0.05
Rain water	0.04	0.01	0.00	0.00	0.01
Tanker truck	0.00	0.00	0.00	0.00	0.00
Water in drums/big cans	0.01	0.00	0.00	0.00	0.00
Bottled water	0.00	0.00	0.00	0.00	0.00
Other	0.02	0.00	0.00	0.00	0.01
Toilet					
Toilet connected to sewer, private	0.08	0.60	0.94	1.00	0.67
Toilet connected to septic well, private	0.44	0.13	0.00	0.00	0.14
Toilet connected to plot/yard, private	0.12	0.03	0.00	0.00	0.04
Traditional pit toilet, private	0.02	0.00	0.00	0.00	0.01
Traditional toilet to sea/river, private	0.02	0.00	0.00	0.00	0.01
Other, private	0.00	0.00	0.00	0.00	0.00
Toilet connected to sewer, shared	0.04	0.23	0.06	0.00	0.08
Toilet connected to septic well, shared	0.05	0.01	0.00	0.00	0.01
Toilet connected to plot/yard, shared	0.01	0.00	0.00	0.00	0.00
Traditional pit toilet, shared	0.00	0.00	0.00	0.00	0.00

Table E-3 PCA factor w eights (example from Colombia, ENDS/ENSIN 2010) (continued)

Variable	Mean within wealth quartile				Total mean
	Q1	Q2	Q3	Q4	
Traditional toilet to sea/river, shared	0.00	0.00	0.00	0.00	0.00
Other, shared	0.00	0.00	0.00	0.00	0.00
No toilet facility	0.20	0.00	0.00	0.00	0.05
Number of toilets	0.8	1.1	1.2	1.7	1.2
Place of toilet					
Outside dwelling area	0.02	0.01	0.00	0.00	0.01
Within plot but outside dwelling	0.43	0.20	0.02	0.00	0.16
Inside dwelling	0.34	0.79	0.97	1.00	0.78
Other	0.00	0.00	0.00	0.00	0.00
Cooking fuel					
Natural gas	0.05	0.30	0.64	0.97	0.50
Propane gas	0.32	0.57	0.28	0.02	0.30
Kerosene/oil/cocinol/diesel/gasoline/alcohol	0.00	0.00	0.00	0.00	0.00
Electricity	0.04	0.05	0.06	0.01	0.04
Firewood, charcoal	0.54	0.02	0.00	0.00	0.13
Mineral coal	0.01	0.00	0.00	0.00	0.00
Do not cook	0.03	0.05	0.01	0.00	0.02
Other	0.00	0.00	0.00	0.00	0.00
Other available services					
Has electricity	0.90	0.99	1.00	1.00	0.97
Access to natural gas	0.06	0.34	0.66	0.98	0.52
Access to piped water	0.55	0.96	1.00	1.00	0.88
Access to sewer	0.14	0.84	1.00	1.00	0.76
Access to garbage collection	0.27	0.94	1.00	1.00	0.81
Landline, exclusive use	0.02	0.16	0.48	0.86	0.38
Landline, shared with others	0.00	0.04	0.03	0.01	0.02
Household items					
Mobile telephone	0.80	0.89	0.94	0.98	0.91
Radio	0.62	0.68	0.83	0.96	0.77
Television	0.72	0.92	0.99	1.00	0.91
Refrigerator	0.42	0.68	0.91	0.99	0.76
Bicycle	0.24	0.31	0.34	0.43	0.33
Motorcycle/scooter	0.13	0.18	0.21	0.19	0.18
Car/truck	0.02	0.05	0.09	0.32	0.12
Shower	0.37	0.83	0.97	1.00	0.80
Canoe	0.03	0.01	0.00	0.00	0.01

Appendix F Estimated parameters from the models with geographic and socioeconomic factors, age, and time as covariates

Table F Estimated parameters from the models with geographic and socioeconomic factors, age, and time as covariates by country and sex

Mexico, men

Variable	μ parameter			σ parameter			ν parameter			τ parameter			Global test for group*
	Coeff.	95% CI	p-value	Coeff.	95% CI	p-value	Coeff.	95% CI	p-value	Coeff.	95% CI	p-value	
Type of residence (ref. urban)													< 0.001
Rural	-0.59	(-0.83, -0.36)	< 0.001	-0.05	(-0.08, -0.01)	< 0.01	-0.01	(-0.23, 0.22)	0.97	-0.01	(-0.12, 0.10)	0.89	
Education (ref. primary)													< 0.01
Secondary	0.24	(0.01, 0.48)	0.04	-0.03	(-0.06, 0.01)	0.19	-0.01	(-0.24, 0.23)	0.94	-0.04	(-0.15, 0.07)	0.49	
Higher	0.42	(0.08, 0.76)	0.02	-0.08	(-0.13, -0.03)	< 0.01	-0.16	(-0.51, 0.19)	0.36	-0.09	(-0.25, 0.08)	0.29	
Household wealth (ref. lowest)													< 0.001
Lower	0.66	(0.38, 0.94)	< 0.001	0.04	(0.00, 0.08)	0.05	0.13	(-0.16, 0.42)	0.39	-0.05	(-0.18, 0.08)	0.44	
Higher	0.74	(0.42, 1.07)	< 0.001	0.05	(0.01, 0.10)	0.03	0.09	(-0.22, 0.39)	0.58	0.05	(-0.11, 0.20)	0.55	
Highest	0.98	(0.61, 1.34)	< 0.001	0.08	(0.02, 0.13)	< 0.01	0.20	(-0.13, 0.53)	0.24	-0.08	(-0.24, 0.07)	0.28	
Age group													< 0.001
20-29	24.8	(24.5, 25.1)	< 0.001	-1.73	(-1.78, -1.69)	< 0.001	-0.69	(-0.99, -0.39)	< 0.001	0.82	(0.67, 0.96)	< 0.001	
30-39	26.9	(26.6, 27.2)	< 0.001	-1.79	(-1.83, -1.75)	< 0.001	-0.43	(-0.73, -0.13)	< 0.01	0.53	(0.38, 0.68)	< 0.001	
40-49	27.4	(27.1, 27.7)	< 0.001	-1.87	(-1.91, -1.82)	< 0.001	-0.26	(-0.56, 0.05)	0.10	0.57	(0.44, 0.70)	< 0.001	
50-59	27.1	(26.8, 27.4)	< 0.001	-1.87	(-1.92, -1.83)	< 0.001	-0.41	(-0.76, -0.07)	0.02	0.50	(0.36, 0.64)	< 0.001	
Type of residence × time (ref. urban)													0.10
Rural	-0.10	(-0.18, -0.02)	0.02	-0.01	(-0.02, 0.00)	0.13	-0.03	(-0.10, 0.05)	0.47	0.00	(-0.04, 0.03)	0.81	
Education × time (ref. primary)													0.47
Secondary	0.07	(0.00, 0.14)	0.07	0.00	(-0.02, 0.01)	0.60	0.05	(-0.03, 0.13)	0.24	-0.02	(-0.05, 0.02)	0.37	
Higher	0.04	(-0.07, 0.15)	0.44	-0.01	(-0.03, 0.00)	0.14	-0.01	(-0.13, 0.11)	0.86	-0.01	(-0.07, 0.04)	0.71	
Household wealth × time (ref. lowest)													0.37
Lower	0.00	(-0.09, 0.08)	0.92	0.01	(-0.01, 0.02)	0.29	-0.02	(-0.11, 0.08)	0.70	0.00	(-0.04, 0.04)	0.85	
Higher	-0.11	(-0.21, -0.01)	0.03	0.01	(0.00, 0.03)	0.19	-0.02	(-0.12, 0.08)	0.70	0.01	(-0.03, 0.06)	0.58	
Highest	-0.03	(-0.14, 0.08)	0.58	0.01	(0.00, 0.03)	0.12	-0.02	(-0.13, 0.09)	0.72	-0.02	(-0.07, 0.03)	0.34	
Age group × time													0.02
20-29	0.06	(-0.03, 0.15)	0.20	0.01	(-0.01, 0.02)	0.27	-0.09	(-0.19, 0.01)	0.07	0.05	(0.01, 0.10)	0.02	
30-39	0.12	(0.03, 0.22)	0.01	0.01	(0.00, 0.02)	0.15	-0.03	(-0.13, 0.07)	0.59	0.01	(-0.03, 0.06)	0.57	
40-49	0.12	(0.02, 0.22)	0.02	0.00	(-0.01, 0.02)	0.94	0.01	(-0.10, 0.11)	0.89	0.02	(-0.02, 0.06)	0.36	
50-59	0.05	(-0.06, 0.16)	0.36	0.00	(-0.02, 0.01)	0.93	-0.02	(-0.13, 0.09)	0.69	-0.03	(-0.08, 0.01)	0.18	

Table F Estimated parameters from the models with geographic and socioeconomic factors, age, and time as covariates by country and sex (continued)

Mexico, women

Variable	μ parameter			σ parameter			ν parameter			τ parameter			Global test for group*
	Coeff.	95% CI	p-value	Coeff.	95% CI	p-value	Coeff.	95% CI	p-value	Coeff.	95% CI	p-value	
Type of residence (ref. urban)													< 0.001
Rural	-0.65	(-0.89, -0.40)	< 0.001	-0.05	(-0.08, -0.02)	< 0.01	-0.03	(-0.22, 0.15)	0.72	0.06	(-0.03, 0.15)	0.18	
Education (ref. primary)													< 0.001
Secondary	-0.63	(-0.87, -0.39)	< 0.001	0.02	(-0.01, 0.05)	0.31	-0.12	(-0.31, 0.07)	0.20	-0.02	(-0.12, 0.07)	0.62	
Higher	-1.83	(-2.22, -1.43)	< 0.001	-0.01	(-0.06, 0.03)	0.61	-0.45	(-0.71, -0.19)	< 0.001	0.08	(-0.07, 0.22)	0.30	
Household wealth (ref. lowest)													< 0.001
Lower	0.52	(0.25, 0.79)	< 0.001	-0.01	(-0.05, 0.03)	0.74	-0.07	(-0.29, 0.15)	0.54	-0.01	(-0.11, 0.10)	0.90	
Higher	0.50	(0.19, 0.81)	< 0.01	-0.02	(-0.06, 0.02)	0.32	-0.12	(-0.36, 0.12)	0.32	0.08	(-0.04, 0.21)	0.20	
Highest	0.09	(-0.25, 0.43)	0.60	-0.05	(-0.09, -0.01)	0.02	-0.24	(-0.49, 0.01)	0.06	0.19	(0.05, 0.33)	< 0.01	
Age group													< 0.001
20-29	26.3	(26.0, 26.6)	< 0.001	-1.60	(-1.65, -1.56)	< 0.001	-0.43	(-0.66, -0.21)	< 0.001	0.62	(0.50, 0.74)	< 0.001	
30-39	28.5	(28.2, 28.8)	< 0.001	-1.67	(-1.71, -1.63)	< 0.001	-0.34	(-0.56, -0.12)	< 0.01	0.53	(0.43, 0.63)	< 0.001	
40-49	29.7	(29.4, 30.0)	< 0.001	-1.70	(-1.74, -1.66)	< 0.001	-0.02	(-0.24, 0.19)	0.84	0.48	(0.38, 0.59)	< 0.001	
50-59	29.9	(29.6, 30.3)	< 0.001	-1.69	(-1.74, -1.65)	< 0.001	-0.11	(-0.35, 0.13)	0.37	0.48	(0.35, 0.60)	< 0.001	
Type of residence × time (ref. urban)													0.67
Rural	-0.03	(-0.11, 0.05)	0.44	0.00	(-0.01, 0.01)	0.41	0.02	(-0.04, 0.08)	0.52	0.00	(-0.03, 0.03)	0.81	
Education × time (ref. primary)													0.99
Secondary	0.03	(-0.04, 0.11)	0.35	0.00	(-0.01, 0.01)	0.94	0.01	(-0.05, 0.07)	0.77	0.01	(-0.02, 0.04)	0.66	
Higher	0.05	(-0.07, 0.17)	0.43	0.00	(-0.01, 0.01)	0.90	0.03	(-0.06, 0.11)	0.56	0.01	(-0.04, 0.05)	0.68	
Household wealth × time (ref. lowest)													0.03
Lower	-0.10	(-0.18, -0.02)	0.01	0.00	(-0.01, 0.01)	0.91	0.07	(0.00, 0.13)	0.05	-0.02	(-0.05, 0.02)	0.33	
Higher	-0.11	(-0.20, -0.01)	0.03	0.00	(-0.01, 0.01)	0.89	0.04	(-0.04, 0.12)	0.32	0.03	(-0.01, 0.07)	0.18	
Highest	-0.11	(-0.21, 0.00)	0.05	-0.01	(-0.02, 0.00)	0.15	0.04	(-0.04, 0.12)	0.38	-0.01	(-0.05, 0.03)	0.67	
Age group × time													0.07
20-29	0.13	(0.04, 0.21)	< 0.01	0.02	(0.01, 0.03)	< 0.01	-0.04	(-0.11, 0.03)	0.24	-0.01	(-0.04, 0.03)	0.59	
30-39	0.12	(0.03, 0.21)	< 0.01	0.01	(0.00, 0.02)	0.13	-0.06	(-0.13, 0.01)	0.11	-0.01	(-0.04, 0.03)	0.66	
40-49	0.17	(0.08, 0.26)	< 0.001	0.00	(-0.01, 0.01)	0.76	-0.01	(-0.08, 0.06)	0.72	-0.02	(-0.05, 0.02)	0.36	
50-59	0.18	(0.08, 0.28)	< 0.001	0.01	(0.00, 0.02)	0.20	-0.06	(-0.14, 0.01)	0.10	0.03	(-0.01, 0.07)	0.18	

Table F Estimated parameters from the models with geographic and socioeconomic factors, age, and time as covariates by country and sex (continued)

Colombia, men

Variable	μ parameter			σ parameter			ν parameter			τ parameter			Global test for group*
	Coeff.	95% CI	p-value	Coeff.	95% CI	p-value	Coeff.	95% CI	p-value	Coeff.	95% CI	p-value	
Type of residence (ref. urban)													< 0.001
Rural	0.19	(0.00, 0.38)	0.05	-0.07	(-0.10, -0.03)	< 0.001	-0.06	(-0.34, 0.22)	0.69	0.01	(-0.09, 0.11)	0.89	
Education (ref. primary)													< 0.001
Secondary	0.51	(0.39, 0.64)	< 0.001	0.02	(-0.01, 0.04)	0.16	0.12	(-0.06, 0.31)	0.18	0.02	(-0.05, 0.09)	0.55	
Higher	0.74	(0.55, 0.92)	< 0.001	0.00	(-0.04, 0.03)	0.79	0.06	(-0.17, 0.30)	0.61	-0.01	(-0.11, 0.08)	0.80	
Household wealth (ref. lowest)													< 0.001
Lower	0.90	(0.71, 1.09)	< 0.001	0.06	(0.03, 0.10)	< 0.001	0.16	(-0.14, 0.46)	0.30	0.07	(-0.03, 0.17)	0.18	
Higher	1.22	(1.00, 1.44)	< 0.001	0.08	(0.03, 0.12)	< 0.001	0.20	(-0.13, 0.52)	0.24	0.00	(-0.13, 0.12)	0.94	
Highest	1.66	(1.43, 1.89)	< 0.001	0.05	(0.01, 0.09)	0.02	0.43	(0.10, 0.77)	0.01	-0.01	(-0.13, 0.11)	0.89	
Age group													< 0.001
20-29	21.9	(21.7, 22.1)	< 0.001	-1.94	(-1.98, -1.90)	< 0.001	-1.21	(-1.52, -0.90)	< 0.001	0.66	(0.55, 0.78)	< 0.001	
30-39	23.9	(23.7, 24.1)	< 0.001	-1.94	(-1.98, -1.91)	< 0.001	-0.68	(-1.01, -0.35)	< 0.001	0.58	(0.47, 0.70)	< 0.001	
40-49	24.3	(24.1, 24.5)	< 0.001	-1.95	(-1.99, -1.91)	< 0.001	-0.46	(-0.77, -0.15)	< 0.01	0.54	(0.43, 0.65)	< 0.001	
50-59	24.4	(24.1, 24.6)	< 0.001	-1.91	(-1.95, -1.87)	< 0.001	-0.34	(-0.65, -0.02)	0.04	0.56	(0.45, 0.68)	< 0.001	
Type of residence × time (ref. urban)													0.47
Rural	-0.03	(-0.09, 0.02)	0.22	0.00	(-0.01, 0.01)	0.42	0.01	(-0.07, 0.08)	0.82	0.01	(-0.02, 0.04)	0.51	
Education × time (ref. primary)													0.02
Secondary	0.06	(0.02, 0.10)	< 0.01	0.00	(-0.01, 0.01)	0.69	-0.01	(-0.07, 0.04)	0.69	0.00	(-0.02, 0.02)	0.92	
Higher	0.05	(-0.01, 0.11)	0.11	0.00	(-0.01, 0.01)	0.52	-0.07	(-0.14, 0.01)	0.08	0.01	(-0.02, 0.04)	0.58	
Household wealth × time (ref. lowest)													0.08
Lower	-0.04	(-0.09, 0.02)	0.17	-0.01	(-0.02, 0.00)	0.10	0.01	(-0.07, 0.09)	0.74	0.00	(-0.02, 0.03)	0.78	
Higher	-0.09	(-0.16, -0.03)	< 0.01	0.00	(-0.01, 0.01)	0.56	-0.02	(-0.11, 0.07)	0.67	0.00	(-0.03, 0.04)	0.86	
Highest	-0.12	(-0.20, -0.05)	< 0.001	-0.01	(-0.02, 0.01)	0.42	-0.01	(-0.11, 0.08)	0.82	0.01	(-0.03, 0.04)	0.67	
Age group × time													0.03
20-29	0.11	(0.05, 0.17)	< 0.001	0.01	(0.00, 0.02)	0.04	0.01	(-0.08, 0.09)	0.88	0.01	(-0.03, 0.04)	0.74	
30-39	0.16	(0.09, 0.22)	< 0.001	0.01	(0.00, 0.02)	0.19	0.02	(-0.07, 0.11)	0.63	-0.02	(-0.05, 0.01)	0.21	
40-49	0.15	(0.09, 0.22)	< 0.001	0.01	(-0.01, 0.02)	0.34	0.02	(-0.07, 0.11)	0.66	-0.04	(-0.07, -0.01)	0.02	
50-59	0.14	(0.07, 0.21)	< 0.001	0.00	(-0.01, 0.01)	0.69	0.01	(-0.09, 0.10)	0.90	-0.01	(-0.04, 0.02)	0.63	

Table F Estimated parameters from the models with geographic and socioeconomic factors, age, and time as covariates by country and sex (continued)

Colombia, women

Variable	μ parameter			σ parameter			ν parameter			τ parameter			Global test for group*
	Coeff.	95% CI	p-value	Coeff.	95% CI	p-value	Coeff.	95% CI	p-value	Coeff.	95% CI	p-value	
Type of residence (ref. urban)													< 0.01
Rural	0.04	(-0.17, 0.25)	0.69	-0.06	(-0.09, -0.03)	< 0.001	-0.11	(-0.30, 0.08)	0.27	-0.01	(-0.09, 0.07)	0.76	
Education (ref. primary)													< 0.001
Secondary	-0.65	(-0.78, -0.52)	< 0.001	-0.03	(-0.05, -0.01)	0.01	-0.21	(-0.35, -0.08)	< 0.01	0.03	(-0.03, 0.09)	0.26	
Higher	-1.25	(-1.42, -1.08)	< 0.001	-0.07	(-0.10, -0.04)	< 0.001	-0.53	(-0.71, -0.35)	< 0.001	0.06	(-0.02, 0.14)	0.16	
Household wealth (ref. lowest)													< 0.001
Lower	0.47	(0.27, 0.67)	< 0.001	-0.05	(-0.08, -0.02)	< 0.001	-0.19	(-0.37, 0.00)	0.04	0.04	(-0.05, 0.12)	0.39	
Higher	0.47	(0.24, 0.70)	< 0.001	-0.10	(-0.14, -0.07)	< 0.001	-0.32	(-0.53, -0.12)	< 0.01	-0.09	(-0.18, 0.00)	0.05	
Highest	0.38	(0.15, 0.61)	< 0.01	-0.17	(-0.21, -0.14)	< 0.001	-0.36	(-0.59, -0.14)	< 0.01	-0.10	(-0.19, 0.00)	0.05	
Age group													< 0.001
20-29	23.9	(23.7, 24.2)	< 0.001	-1.66	(-1.69, -1.63)	< 0.001	-0.32	(-0.53, -0.11)	< 0.01	0.61	(0.51, 0.71)	< 0.001	
30-39	25.9	(25.7, 26.1)	< 0.001	-1.66	(-1.69, -1.63)	< 0.001	-0.15	(-0.36, 0.05)	0.14	0.62	(0.53, 0.71)	< 0.001	
40-49	27.0	(26.8, 27.3)	< 0.001	-1.64	(-1.67, -1.61)	< 0.001	0.07	(-0.13, 0.26)	0.49	0.58	(0.48, 0.67)	< 0.001	
50-59	27.8	(27.5, 28.0)	< 0.001	-1.62	(-1.66, -1.59)	< 0.001	0.29	(0.07, 0.52)	< 0.01	0.57	(0.47, 0.67)	< 0.001	
Type of residence × time (ref. urban)													0.65
Rural	-0.04	(-0.10, 0.01)	0.12	0.00	(-0.01, 0.01)	0.76	-0.01	(-0.06, 0.04)	0.67	0.00	(-0.02, 0.03)	0.81	
Education × time (ref. primary)													0.39
Secondary	0.01	(-0.03, 0.05)	0.61	0.00	(-0.01, 0.01)	0.75	-0.02	(-0.06, 0.02)	0.31	0.01	(-0.01, 0.03)	0.24	
Higher	0.04	(-0.01, 0.09)	0.10	0.00	(-0.01, 0.01)	1.00	-0.03	(-0.09, 0.02)	0.25	0.02	(-0.01, 0.04)	0.15	
Household wealth × time (ref. lowest)													< 0.001
Lower	-0.08	(-0.14, -0.03)	< 0.01	0.00	(-0.01, 0.01)	0.79	-0.07	(-0.12, -0.01)	0.02	0.02	(0.00, 0.04)	0.08	
Higher	-0.12	(-0.19, -0.06)	< 0.001	-0.01	(-0.02, 0.00)	0.15	-0.04	(-0.10, 0.01)	0.14	0.00	(-0.03, 0.02)	0.90	
Highest	-0.14	(-0.21, -0.07)	< 0.001	-0.01	(-0.02, 0.00)	0.08	0.00	(-0.06, 0.07)	0.91	0.00	(-0.03, 0.03)	0.86	
Age group × time													0.02
20-29	0.20	(0.14, 0.26)	< 0.001	0.02	(0.01, 0.03)	< 0.01	0.06	(0.00, 0.12)	0.04	-0.02	(-0.05, 0.01)	0.19	
30-39	0.22	(0.16, 0.28)	< 0.001	0.01	(0.00, 0.02)	0.16	0.07	(0.01, 0.13)	0.02	-0.03	(-0.05, 0.00)	0.05	
40-49	0.20	(0.14, 0.26)	< 0.001	0.00	(-0.01, 0.01)	0.47	0.03	(-0.03, 0.09)	0.30	-0.02	(-0.05, 0.00)	0.11	
50-59	0.18	(0.11, 0.25)	< 0.001	0.00	(-0.01, 0.01)	0.78	0.04	(-0.02, 0.10)	0.23	-0.02	(-0.05, 0.01)	0.17	

Table F Estimated parameters from the models with geographic and socioeconomic factors, age, and time as covariates by country and sex (continued)

Peru, men

Variable	μ parameter			σ parameter			ν parameter			τ parameter			Global test for group*
	Coeff.	95% CI	p-value	Coeff.	95% CI	p-value	Coeff.	95% CI	p-value	Coeff.	95% CI	p-value	
Type of residence (ref. urban)													< 0.001
Rural	-0.58	(-0.80, -0.35)	< 0.001	-0.06	(-0.11, -0.02)	0.01	-0.08	(-0.46, 0.30)	0.69	-0.02	(-0.15, 0.12)	0.83	
Education (ref. primary)													< 0.001
Secondary	0.46	(0.29, 0.63)	< 0.001	0.02	(-0.03, 0.06)	0.48	-0.17	(-0.56, 0.23)	0.41	-0.01	(-0.15, 0.12)	0.84	
Higher	0.30	(0.07, 0.53)	0.01	-0.03	(-0.09, 0.02)	0.23	-0.41	(-0.87, 0.05)	0.08	-0.04	(-0.20, 0.12)	0.62	
Household wealth (ref. lowest)													< 0.001
Lower	1.10	(0.89, 1.32)	< 0.001	0.16	(0.11, 0.22)	< 0.001	-0.39	(-0.88, 0.10)	0.12	0.06	(-0.09, 0.21)	0.45	
Higher	1.81	(1.51, 2.11)	< 0.001	0.22	(0.16, 0.29)	< 0.001	-0.14	(-0.71, 0.43)	0.63	0.00	(-0.19, 0.18)	0.98	
Highest	2.56	(2.21, 2.92)	< 0.001	0.26	(0.18, 0.33)	< 0.001	0.31	(-0.29, 0.91)	0.31	-0.06	(-0.27, 0.15)	0.59	
Age group													< 0.001
20-29	22.4	(22.1, 22.7)	< 0.001	-2.17	(-2.24, -2.10)	< 0.001	0.00	(-0.54, 0.54)	1.00	0.53	(0.35, 0.72)	< 0.001	
30-39	23.8	(23.6, 24.1)	< 0.001	-2.19	(-2.25, -2.12)	< 0.001	-0.27	(-0.79, 0.26)	0.32	0.52	(0.34, 0.70)	< 0.001	
40-49	24.3	(24.0, 24.5)	< 0.001	-2.18	(-2.24, -2.11)	< 0.001	-0.08	(-0.62, 0.46)	0.78	0.61	(0.44, 0.79)	< 0.001	
50-59	24.3	(24.0, 24.6)	< 0.001	-2.14	(-2.21, -2.07)	< 0.001	0.04	(-0.51, 0.60)	0.88	0.57	(0.39, 0.75)	< 0.001	
Type of residence × time (ref. urban)													0.20
Rural	0.03	(-0.07, 0.13)	0.55	-0.02	(-0.04, 0.01)	0.15	-0.06	(-0.23, 0.10)	0.45	0.06	(-0.01, 0.12)	0.07	
Education × time (ref. primary)													0.75
Secondary	0.03	(-0.05, 0.10)	0.44	0.00	(-0.02, 0.02)	0.90	-0.07	(-0.25, 0.10)	0.40	-0.02	(-0.08, 0.04)	0.46	
Higher	0.00	(-0.10, 0.11)	0.94	-0.02	(-0.04, 0.01)	0.22	-0.10	(-0.29, 0.10)	0.33	-0.01	(-0.09, 0.06)	0.74	
Household wealth × time (ref. lowest)													< 0.01
Lower	0.17	(0.07, 0.27)	< 0.001	-0.01	(-0.03, 0.01)	0.45	0.15	(-0.07, 0.37)	0.18	0.03	(-0.04, 0.10)	0.37	
Higher	0.04	(-0.09, 0.18)	0.52	-0.03	(-0.06, 0.00)	0.03	0.24	(-0.01, 0.49)	0.06	-0.02	(-0.10, 0.06)	0.59	
Highest	0.11	(-0.04, 0.27)	0.15	-0.02	(-0.05, 0.02)	0.32	0.27	(0.01, 0.53)	0.04	0.00	(-0.09, 0.09)	0.96	
Age group × time													0.50
20-29	0.04	(-0.08, 0.15)	0.53	0.03	(0.00, 0.06)	0.03	-0.15	(-0.39, 0.10)	0.24	0.02	(-0.06, 0.10)	0.69	
30-39	0.16	(0.04, 0.27)	< 0.01	0.04	(0.01, 0.07)	< 0.01	-0.10	(-0.33, 0.13)	0.40	0.02	(-0.07, 0.10)	0.71	
40-49	0.14	(0.02, 0.26)	0.02	0.03	(0.00, 0.06)	0.03	-0.02	(-0.26, 0.22)	0.89	0.01	(-0.07, 0.09)	0.74	
50-59	0.12	(-0.01, 0.24)	0.07	0.03	(0.00, 0.06)	0.04	-0.12	(-0.36, 0.12)	0.34	-0.02	(-0.10, 0.06)	0.67	

Table F Estimated parameters from the models with geographic and socioeconomic factors, age, and time as covariates by country and sex (continued)

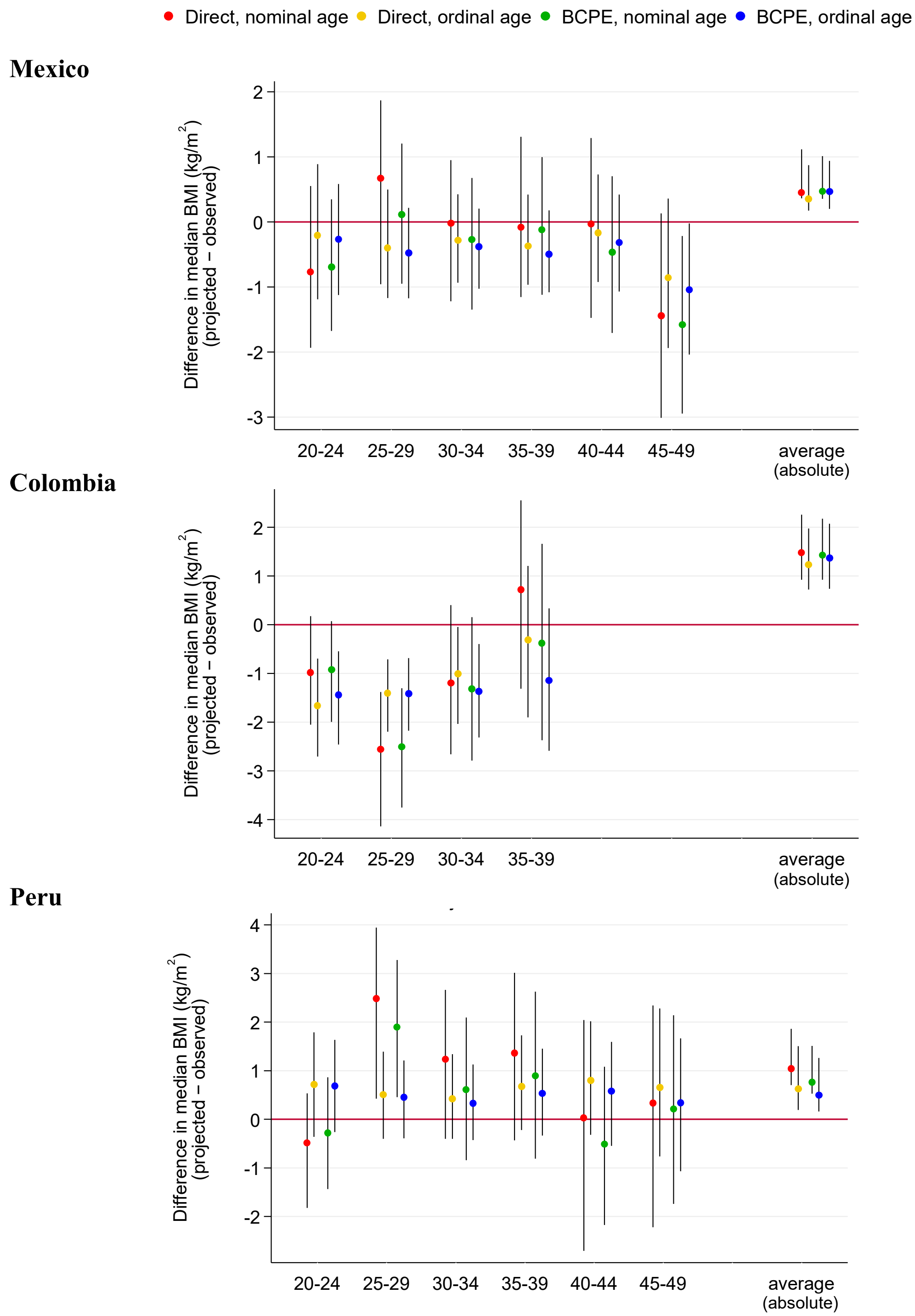
Peru, women

Variable	μ parameter			σ parameter			ν parameter			τ parameter			Global test for group*
	Coeff.	95% CI	p-value	Coeff.	95% CI	p-value	Coeff.	95% CI	p-value	Coeff.	95% CI	p-value	
Type of residence (ref. urban)													0.24
Rural	-0.21	(-0.46, 0.04)	0.11	-0.04	(-0.08, 0.01)	0.12	0.01	(-0.29, 0.31)	0.95	0.00	(-0.13, 0.13)	0.99	
Education (ref. primary)													< 0.001
Secondary	-0.18	(-0.40, 0.04)	0.10	-0.03	(-0.07, 0.02)	0.20	-0.25	(-0.56, 0.06)	0.11	-0.07	(-0.20, 0.06)	0.27	
Higher	-1.21	(-1.48, -0.94)	< 0.001	-0.11	(-0.16, -0.06)	< 0.001	-0.52	(-0.90, -0.14)	< 0.01	-0.02	(-0.18, 0.14)	0.81	
Household wealth (ref. lowest)													< 0.001
Lower	1.81	(1.55, 2.07)	< 0.001	0.08	(0.03, 0.12)	< 0.01	0.14	(-0.20, 0.48)	0.42	-0.10	(-0.24, 0.03)	0.13	
Higher	2.28	(1.93, 2.62)	< 0.001	0.08	(0.02, 0.14)	0.01	0.05	(-0.36, 0.46)	0.81	-0.05	(-0.22, 0.12)	0.56	
Highest	2.18	(1.79, 2.57)	< 0.001	0.10	(0.02, 0.17)	< 0.01	0.02	(-0.45, 0.49)	0.94	-0.19	(-0.38, 0.01)	0.06	
Age group													< 0.001
20-29	22.9	(22.6, 23.2)	< 0.001	-1.92	(-1.98, -1.86)	< 0.001	-0.67	(-1.08, -0.26)	< 0.01	0.74	(0.56, 0.92)	< 0.001	
30-39	25.0	(24.7, 25.3)	< 0.001	-1.88	(-1.94, -1.83)	< 0.001	-0.15	(-0.57, 0.27)	0.48	0.72	(0.55, 0.89)	< 0.001	
40-49	25.6	(25.3, 25.9)	< 0.001	-1.91	(-1.96, -1.85)	< 0.001	-0.35	(-0.72, 0.02)	0.07	0.68	(0.52, 0.85)	< 0.001	
50-59	25.7	(25.3, 26.0)	< 0.001	-1.80	(-1.86, -1.74)	< 0.001	-0.04	(-0.43, 0.35)	0.86	0.65	(0.49, 0.82)	< 0.001	
Type of residence × time (ref. urban)													0.22
Rural	-0.05	(-0.17, 0.06)	0.35	-0.01	(-0.02, 0.01)	0.60	-0.15	(-0.29, -0.02)	0.03	0.03	(-0.03, 0.08)	0.34	
Education × time (ref. primary)													0.11
Secondary	-0.08	(-0.18, 0.01)	0.09	0.01	(-0.01, 0.03)	0.52	-0.03	(-0.16, 0.11)	0.69	0.04	(-0.02, 0.09)	0.22	
Higher	-0.08	(-0.20, 0.03)	0.16	-0.01	(-0.03, 0.02)	0.53	0.10	(-0.07, 0.27)	0.24	0.00	(-0.08, 0.07)	0.93	
Household wealth × time (ref. lowest)													0.10
Lower	-0.01	(-0.13, 0.10)	0.81	0.00	(-0.02, 0.02)	0.96	-0.15	(-0.30, 0.00)	0.05	0.00	(-0.05, 0.06)	0.89	
Higher	-0.09	(-0.23, 0.06)	0.23	-0.01	(-0.04, 0.02)	0.43	-0.07	(-0.25, 0.11)	0.44	0.01	(-0.07, 0.08)	0.88	
Highest	-0.20	(-0.36, -0.03)	0.02	0.00	(-0.03, 0.04)	0.81	-0.10	(-0.31, 0.11)	0.35	-0.05	(-0.14, 0.03)	0.21	
Age group × time													0.02
20-29	0.25	(0.11, 0.39)	< 0.001	0.02	(-0.01, 0.05)	0.12	0.03	(-0.16, 0.21)	0.79	0.01	(-0.07, 0.09)	0.83	
30-39	0.37	(0.24, 0.51)	< 0.001	0.01	(-0.02, 0.03)	0.60	0.15	(-0.04, 0.33)	0.12	-0.03	(-0.10, 0.05)	0.45	
40-49	0.43	(0.29, 0.56)	< 0.001	0.02	(-0.01, 0.04)	0.22	0.19	(0.02, 0.35)	0.03	-0.03	(-0.10, 0.05)	0.48	
50-59	0.38	(0.23, 0.53)	< 0.001	0.01	(-0.02, 0.03)	0.66	0.26	(0.09, 0.43)	< 0.01	-0.04	(-0.11, 0.03)	0.25	

* P-values are from $4(k_j - 1)$ *df* likelihood ratio tests. The null hypothesis is: $\beta^{\mu}_{mjk} = 0, \beta^{\sigma}_{mjk} = 0, \beta^{\nu}_{mjk} = 0, \beta^{\tau}_{mjk} = 0$, $k = 2, 3, \dots, k_j$, $m = 1$ or 2, where β^{i}_{mjk} is the coefficient for the category k of the factor $j =$ (place of residence, education, household wealth, age) (when $m = 1$) or for the interaction term between the category k of the factor j and time (when $m = 2$) for the parameter $i = (\mu, \sigma, \nu, \tau)$.

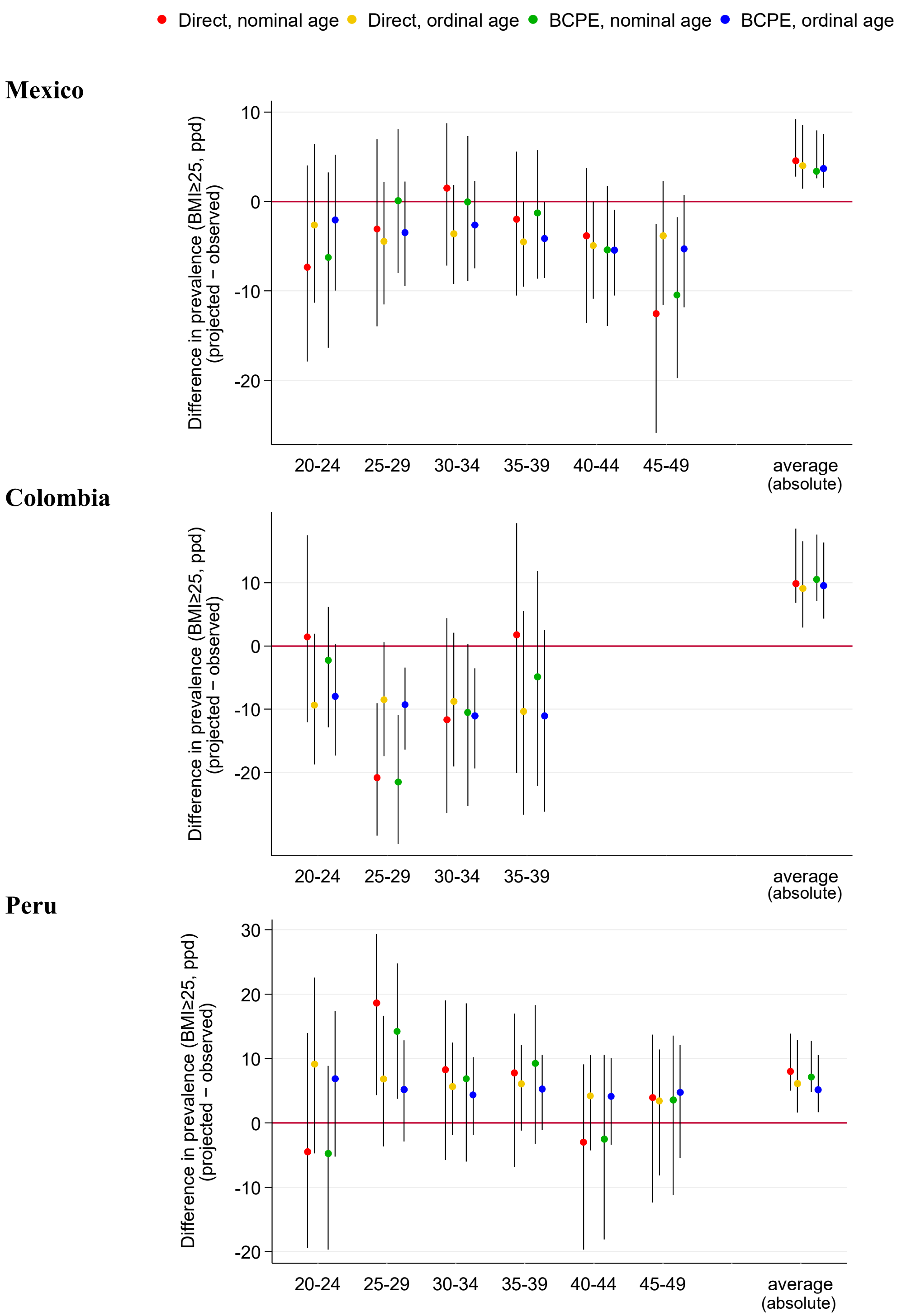
Appendix G Prediction errors by age

Figure G-1 Prediction errors for median BMI by age



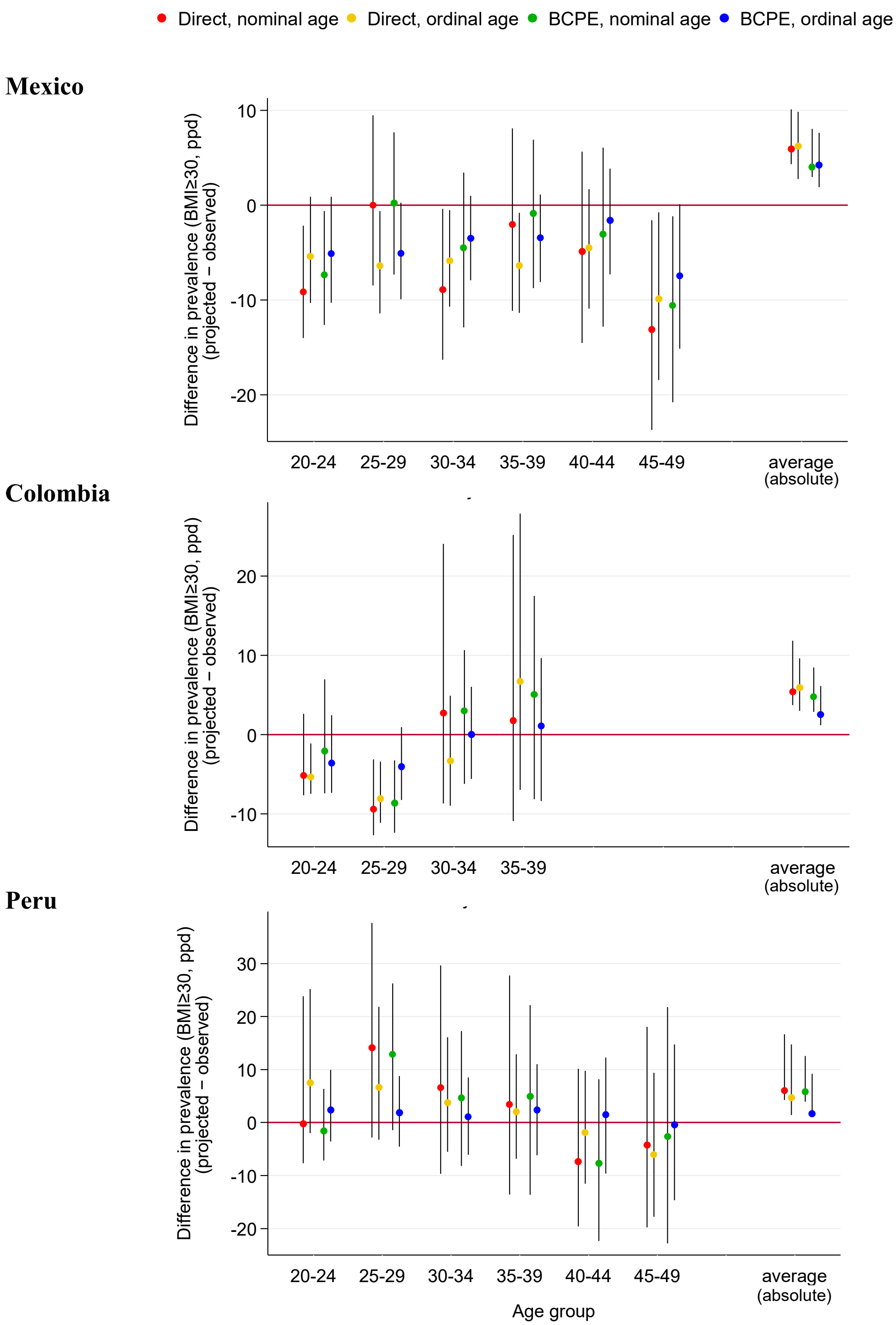
Direct, direct method; BCPE, BCPE method; nominal age, 5-year age group included as a nominal variable; ordinal age, 5-year age group included as an ordinal variable.

Figure G-2 Prediction errors for overweight and obesity prevalence (BMI ≥ 25 kg/m²) by age



Direct, direct method; BCPE, BCPE method; nominal age, 5-year age group included as a nominal variable; ordinal age, 5-year age group included as an ordinal variable.

Figure G-3 Prediction errors for obesity prevalence (BMI ≥ 30 kg/m^2) by age



Direct, direct method; BCPE, BCPE method; nominal age, 5-year age group included as a nominal variable; ordinal age, 5-year age group included as an ordinal variable.

Bibliography

Aitsi-Selmi, A., Bell, R., Shipley, M. J., & Marmot, M. G. (2014). Education modifies the association of wealth with obesity in women in middle-income but not low-income countries: An interaction study using seven national datasets, 2005-2010. *PloS One, 9*(3), e90403.

Alvarez-Dongo, D., Sanchez-Abanto, J., Gomez-Guizado, G., & Tarqui-Mamani, C. (2012). Overweight and obesity: Prevalence and determining social factors of overweight in the Peruvian population (2009-2010). *Revista Peruana de Medicina Experimental y Salud Publica, 29*(3), 303-313.

Arterburn, D. E., Crane, P. K., & Sullivan, S. D. (2004). The coming epidemic of obesity in elderly Americans. *Journal of the American Geriatrics Society, 52*(11), 1907-1912.

Barendregt, J. J., & Veerman, J. L. (2010). Categorical versus continuous risk factors and the calculation of potential impact fractions. *Journal of Epidemiology and Community Health, 64*(3), 209-212.

Basu, A. (2010). Forecasting distribution of body mass index in the United States: Is there more room for growth? *Medical Decision Making: An International Journal of the Society for Medical Decision Making, 30*(3), E1-E11.

Beran, D. (2015). The impact of health systems on diabetes care in low and lower middle income countries. *Current Diabetes Reports, 15*(4), 20-015-0591-8.

Bray, G. A. (2004). Medical consequences of obesity. *The Journal of Clinical Endocrinology and Metabolism, 89*(6), 2583-2589.

Caballero, B. (2007). The global epidemic of obesity: An overview. *Epidemiologic Reviews, 29*, 1-5.

Claro, R. M., Levy, R. B., Popkin, B. M., & Monteiro, C. A. (2012). Sugar-sweetened beverage taxes in Brazil. *American Journal of Public Health*, 102(1), 178-183.

Centro Nacional de Alimentación y Nutrición (CENAN), & Instituto Nacional de Estadística e Informática (INEI). (2009). *Informe final: Perfil nutricional y pobreza en Perú. ENAHO I trimestre 2008*. Lima, Peru: CENAN and INEI.

Chan, J. S., Yan, J. H., & Payne, V. G. (2013). The impact of obesity and exercise on cognitive aging. *Frontiers in Aging Neuroscience, 5*, 97.

Colchero, M. A., Salgado, J. C., Unar-Munguia, M., Hernandez-Avila, M., & Rivera-Dommarco, J. A. (2015). Price elasticity of the demand for sugar sweetened beverages and soft drinks in Mexico. *Economics and Human Biology*, 19, 129-137.

Corsi, D. J., Finlay, J. E., & Subramanian, S. V. (2012). Weight of communities: A multilevel analysis of body mass index in 32,814 neighborhoods in 57 low- to middle-income countries (LMICs). *Social Science & Medicine, 75*(2), 311-322.

Cuevas, A., Alvarez, V., & Olivos, C. (2009). The emerging obesity problem in Latin America. *Expert Review of Cardiovascular Therapy, 7*(3), 281-288.

Dinsa, G. D., Goryakin, Y., Fumagalli, E., & Suhrcke, M. (2012). Obesity and socioeconomic status in developing countries: A systematic review. *Obesity Reviews, 13*(11), 1067-1079.

Dirección Ejecutiva de vigilancia Alimentaria y Nutricional (DEVAN). (2015b). *Encuesta nacional de hogar - modulo mediciones antropométricas, 2012-2013 [datasets]*. Lima, Peru: DEVAN [producer]. Instituto Nacional de Salud (INS) [distributor]. http://www.portal.ins.gob.pe/es/cenan/cenan-c2/vigilancia-alimentaria-y-nutricional/vigilancia-de-indicadores-nutricionales-b.

Dirección Ejecutiva de vigilancia Alimentaria y Nutricional (DEVAN). (2015a). *Informe técnico: Estado nutricional en el Perú por etapas de vida; 2012-2013*. Lima, Peru: DEVAN.

Egger, G., Swinburn, B., & Islam, F. M. (2012). Economic growth and obesity: An interesting relationship with world-wide implications. *Economics and Human Biology, 10*(2), 147-153.

Emerging Risk Factors Collaboration, Wormser, D., Kaptoge, S., Di Angelantonio, E., Wood, A. M., Pennells, L. et al. (2011). Separate and combined associations of body-mass index and abdominal adiposity with cardiovascular disease: Collaborative analysis of 58 prospective studies. *Lancet, 377*(9771), 1085-1095.

Ezzati, M., Vander Hoorn, S., Lawes, C. M., Leach, R., James, W. P., Lopez, A. D. et al. (2005). Rethinking the "diseases of affluence" paradigm: Global patterns of nutritional risks in relation to economic development. *PLoS Medicine, 2*(5), e133.

Finkelstein, E. A., Ruhm, C. J., & Kosa, K. M. (2005). Economic causes and consequences of obesity. *Annual Review of Public Health, 26*, 239-257.

Finucane, M. M., Stevens, G. A., Cowan, M. J., Danaei, G., Lin, J. K., Paciorek, C. J. et al. (2011). National, regional, and global trends in body-mass index since 1980: Systematic analysis of health examination surveys and epidemiological studies with 960 country-years and 9.1 million participants. *Lancet (London, England), 377*(9765), 557-567.

Flegal, K. M., Carroll, M. D., Kit, B. K., & Ogden, C. L. (2012). Prevalence of obesity and trends in the distribution of body mass index among US adults, 1999-2010. *JAMA, 307*(5), 491-497.

Fleischer, N. L., Diez Roux, A. V., & Hubbard, A. E. (2012). Inequalities in body mass index and smoking behavior in 70 countries: Evidence for a social transition in chronic disease risk. *American Journal of Epidemiology, 175*(3), 167-176.

Fraser, B. (2005). Latin America's urbanisation is boosting obesity. *Lancet (London, England), 365*(9476), 1995-1996.

Fry, K., Firestone, R., & Chakraborty, N. M. (2014). *Measuring equity with nationally representative wealth quintiles*. Washington, DC, USA: PSI.

Galobardes, B., Shaw, M., Lawlor, D. A., Lynch, J. W., & Davey Smith, G. (2006). Indicators of socioeconomic position (part 1). *Journal of Epidemiology and Community Health, 60*(1), 7-12.

Goryakin, Y., Lobstein, T., James, W. P., & Suhrcke, M. (2015). The impact of economic, political and social globalization on overweight and obesity in the 56 low and middle income countries. *Social Science & Medicine, 133*, 67-76.

Goryakin, Y., & Suhrcke, M. (2014). Economic development, urbanization, technological change and overweight: What do we learn from 244 demographic and health surveys? *Economics and Human Biology, 14*, 109-127.

Gutiérrez, J. P., Rivera-Dommarco, J., Shamah-Levy, T., Villalpando-Hernández, S., Franco, A., Cuevas-Nasu, L. et al. (2013). *Encuesta nacional de salud y nutrición 2012. resultados nacionales.* (2nd ed.). Cuernavaca, Mexico: Instituto Nacional de Salud Pública (INSP).

Hill, J. O., Wyatt, H. R., Reed, G. W., & Peters, J. C. (2003). Obesity and the environment: Where do we go from here? *Science (New York, N.Y.), 299*(5608), 853-855.

Howel, D. (2011). Trends in the prevalence of obesity and overweight in English adults by age and birth cohort, 1991-2006. *Public Health Nutrition, 14*(1), 27-33.

Institute for Health Metrics and Evaluation. (n.d.). *Overweight and obesity viz.* Retrieved 4/8, 2016 from http://vizhub.healthdata.org/obesity

Instituto Colombiano de Bienestar Familiar (ICBF). (2011). *Encuesta nacional de la situación nutricional en Colombia 2010 - ENSIN*. Bogotá, Colombia: ICBF.

Instituto Colombiano de Bienestar Familiar (ICBF). (2006). *Encuesta nacional de la situación nutricional en Colombia, 2005*. Bogotá, Colombia: ICBF.

Instituto Nacional de Estadística e Informática (INEI). (2015b). *Encuesta demográfica y de salud familiar, ENDES continua, 2014 [datasets]*. Lima, Peru: INEI [producer]. Rockville, USA: The DHS Program/ICF International [distributor]. https://dhsprogram.com/data.

Instituto Nacional de Estadística e Informática (INEI). (2015a). *Perú: Encuesta demográfica y de salud familiar - ENDES 2014, nacional y departamental*. Lima, Peru: INEI.

Instituto Nacional de Estadística e Informática (INEI). (2014c). *Encuesta nacional de hogares 2013: Metodología del cálculo del factor de expansión*. Lima, Peru: INEI.

Instituto Nacional de Estadística e Informática (INEI). (2014b). *Encuesta demográfica y de salud familiar, ENDES continua, 2013 [datasets]*. Lima, Peru: INEI [producer]. Rockville, USA: The DHS Program/ICF International [distributor]. https://dhsprogram.com/data.

Instituto Nacional de Estadística e Informática (INEI). (2014a). *Perú: Encuesta demográfica y de salud familiar - ENDES 2013, nacional y departamental*. Lima, Peru: INEI.

Instituto Nacional de Estadística e Informática (INEI). (2013d). *Encuesta demográfica y de salud familiar, ENDES continua, 2012 [datasets]*. Lima, Peru: INEI [producer]. Rockville, USA: The DHS Program/ICF International [distributor]. https://dhsprogram.com/data.

Instituto Nacional de Estadística e Informática (INEI). (2013c). *Perú: Encuesta demográfica y de salud familiar 2012, nacional y departamental*. Lima, Peru: INEI.

Instituto Nacional de Estadística e Informática (INEI). (2013b). *Encuesta nacional de hogares sobre condiciones de vida y pobreza, ENAHO 2012 [datasets]*. Lima, Peru: INEI [producer and distributor]. http://iinei.inei.gob.pe/microdatos.

Instituto Nacional de Estadística e Informática (INEI). (2013a). *Ficha técnica: Encuesta nacional de hogares sobre condiciones de vida y pobreza - ENAHO 2012*. Lima, Peru: INEI.

Instituto Nacional de Estadística e Informática (INEI). (2012f). *Encuesta demográfica y de salud familiar, ENDES continua 2011*. Lima, Peru: INEI [producer]. Rockville, USA: The DHS Program/ICF International [distributor]. https://dhsprogram.com/data.

Instituto Nacional de Estadística e Informática (INEI). (2012e). *Perú: Encuesta demográfica y de salud familiar 2011, nacional y departamental*. Lima, Peru: INEI.

Instituto Nacional de Estadística e Informática (INEI). (2012d). *Encuesta nacional de hogares sobre condiciones de vida y pobreza, ENAHO 2008 [datasets]*. Lima, Peru: INEI [producer and distributor]. http://iinei.inei.gob.pe/microdatos.

Instituto Nacional de Estadística e Informática (INEI). (2012c). *Encuesta nacional de hogares sobre condiciones de vida y pobreza, ENAHO 2007 [datasets]*. Lima, Peru: INEI [producer and distributor]. http://iinei.inei.gob.pe/microdatos.

Instituto Nacional de Estadística e Informática (INEI). (2012b). *Ficha técnica: Encuesta nacional de hogares sobre condiciones de vida y pobreza - ENAHO 2008*. Lima, Peru: INEI.

Instituto Nacional de Estadística e Informática (INEI). (2012a). *Ficha técnica: Encuesta nacional de hogares sobre condiciones de vida y pobreza - ENAHO 2007*. Lima, Peru: INEI.

Instituto Nacional de Estadística e Informática (INEI), & Centro Nacional de Alimentación y Nutrición (CENAN). (2014). *Monitoreo de indicadores nutricionales en la encuesta nacional de hogares (ENAHO) 2011: Diccionario de datos*. Lima, Peru: INEI-CENAN.

Instituto Nacional de Estadística e Informática (INEI), & Centro Nacional de Alimentación y Nutrición (CENAN). (2012). *Monitoreo de indicadores nutricionales en la ENAHO 2009-2010: Diccionario de datos*. Lima, Peru: INEI-CENAN.

Instituto Nacional de Estadística e Informática (INEI), & Centro Nacional de Alimentación y Nutrición (CENAN). (2010b). *Encuesta nacional de hogares - modulo mediciones antropométricas, 2007-2008 [datasets]*. Lima, Peru: INEI and CENAN [producer]. INEI [distributor]. http://iinei.inei.gob.pe/microdatos.

Instituto Nacional de Estadística e Informática (INEI), & Centro Nacional de Alimentación y Nutrición (CENAN). (2010a). *Monitoreo de indicadores nutricionales en la ENAHO 2007-2008: Diccionario de datos*. Lima, Peru: INEI-CENAN.

Instituto Nacional de Estadística e Informática (INEI), & Macro International. (2001b). *Encuesta demográfica y de salud familiar, ENDES, 2000 [datasets]*. Lima, Peru: INEI, and Calverton, USA: Macro International [producer]. Rockville, USA: The DHS Program/ICF International [distributor]. https://dhsprogram.com/data.

Instituto Nacional de Estadística e Informática (INEI), & Macro International. (2001a). *Perú: Encuesta demográfica y de salud familiar 2000*. Lima, Peru: INEI, and Calverton, USA: Macro International.

Instituto Nacional de Estadística e Informática (INEI), & ORC Macro. (2011b). *Encuesta demográfica y de salud familiar, ENDES continua, 2010 [datasets]*. Lima, Peru: INEI, and Calverton, USA: ORC Macro [producer]. Rockville, USA: The DHS Program/ICF International [distributor]. https://dhsprogram.com/data.

Instituto Nacional de Estadística e Informática (INEI), & ORC Macro. (2011a). *Perú: Encuesta demográfica y de salud familiar ENDES continua, 2010, informe principal*. Lima, Peru: INEI, and Calverton, USA: ORC Macro.

Instituto Nacional de Estadística e Informática (INEI), & ORC Macro. (2010c). *Encuesta demográfica y de salud familiar, ENDES continua, 2009 [datasets]*. Lima, Peru: INEI, and Calverton, USA: ORC Macro [producer]. Rockville, USA: The DHS Program/ICF International [distributor]. https://dhsprogram.com/data.

Instituto Nacional de Estadística e Informática (INEI), & ORC Macro. (2010b). *Encuesta demográfica y de salud familiar, ENDES continua, 2004-2008 [datasets]*. Lima, Peru: INEI, and Calverton, USA: ORC Macro [producer]. Rockville, USA: The DHS Program/ICF International [distributor]. https://dhsprogram.com/data.

Instituto Nacional de Estadística e Informática (INEI), & ORC Macro. (2010a). *Perú: Encuesta demográfica y de salud familiar ENDES continua, 2009, informe principal*. Lima, Peru: INEI and Calverton, USA: ORC Macro.

Instituto Nacional de Estadística e Informática (INEI), & ORC Macro. (2009). *Perú: Encuesta demográfica y de salud familiar 2007-2008, informe principal*. Lima, Peru: INEI, and Calverton, USA: ORC Macro.

Instituto Nacional de Estadística e Informática (INEI), & ORC Macro. (2007). *Perú: Encuesta demográfica y de salud familiar 2004-2006*. Lima, Peru: INEI, and Calverton, USA: ORC Macro.

Instituto Nacional de Salud Pública (INSP). (2012b). *Encuesta nacional de salud y nutrición (ENSANUT) 2012 [datasets]*. Cuernavaca, Mexico: INSP [producer and distributor]. http://ensanut.insp.mx.

Instituto Nacional de Salud Pública (INSP). (2012a). *Encuesta nacional de salud y nutrición (ENSANUT) 2006 [datasets]*. Cuernavaca, Mexico: INSP [producer and distributor]. http://ensanut.insp.mx.

Instituto Nacional de Salud Pública (INSP). (2009b). *Encuesta nacional de salud (ENSA) 2000 [datasets]*. Cuernavaca, Mexico: INSP [producer and distributor]. http://www.insp.mx/encuestoteca.html.

Instituto Nacional de Salud Pública (INSP). (2009a). *Encuesta nacional de nutrición (ENN) 1988 [datasets]*. Cuernavaca, Mexico: INSP [producer and distributor]. http://www.insp.mx/encuestoteca.html.

International Diabetes Federation (IDF). (2015). *IDF diabetes atlas* (7th ed.). Brussels, Belgium: IDF.

Jacoby, E., Goldstein, J., Lopez, A., Nunez, E., & Lopez, T. (2003). Social class, family, and life-style factors associated with overweight and obesity among adults in Peruvian cities. *Preventive Medicine, 37*(5), 396-405.

James, W. P. T. (2008). WHO recognition of the global obesity epidemic. *International Journal Obesity (London), 32*(Suppl 7), S120-S126.

Jenkins, K. R., & Fultz, N. H. (2008). The relationship of older adults' activities and body mass index. *Journal of Aging and Health, 20*(2), 217-234.

Jones-Smith, J. C., Gordon-Larsen, P., Siddiqi, A., & Popkin, B. M. (2012). Is the burden of overweight shifting to the poor across the globe? time trends among women in 39 low- and middle-income countries (1991-2008). *International Journal of Obesity (2005), 36*(8), 1114-1120.

Jones-Smith, J. C., Gordon-Larsen, P., Siddiqi, A., & Popkin, B. M. (2011). Cross-national comparisons of time trends in overweight inequality by socioeconomic status among women using repeated cross-sectional surveys from 37 developing countries, 1989-2007. *American Journal of Epidemiology, 173*(6), 667-675.

Kanter, R., & Caballero, B. (2012). Global gender disparities in obesity: A review. *Advances in Nutrition (Bethesda, Md.), 3*(4), 491-498.

Kasper, N. M., Herran, O. F., & Villamor, E. (2014). Obesity prevalence in Colombian adults is increasing fastest in lower socio-economic status groups and urban residents: Results from two nationally representative surveys. *Public Health Nutrition, 17*(11), 2398-2406.

Koenker, R., & Bassett, G. (1978). Regression quantiles. *Journal of Economic Perspectives, 46*(1), 33-50.

Kolenikov, S. (2010). Resampling variance estimation for complex survey data. *Stata Journal, 10*(2), 165-199.

Kopelman, P. (2007). Health risks associated with overweight and obesity. *Obesity Reviews, 8 Suppl 1*, 13-17.

Krishna, A., Razak, F., Lebel, A., Smith, G. D., & Subramanian, S. V. (2015). Trends in group inequalities and interindividual inequalities in BMI in the United States, 1993-2012. *The American Journal of Clinical Nutrition, 101*(3), 598-605.

Labarthe, D. (2011). Chapter 10. obesity. *Epidemiology and prevention of cardiovascular diseases: A global challenge* (2nd ed.) (pp. 223). Sudbury, Massachusetts, USA: Jones and Bartlett Publishers.

Lanas, F., Avezum, A., Bautista, L. E., Diaz, R., Luna, M., Islam, S. et al. (2007). Risk factors for acute myocardial infarction in Latin America: The INTERHEART Latin American study. *Circulation, 115*(9), 1067-1074.

Lanata, C. F. (2012). Overweight and obesity in peru: Urgent need to have a public health policy for their control. *Revista Peruana de Medicina Experimental y Salud Publica, 29*(3), 299-300.

Lee, R. D., & Carter, L. R. (1992). Modeling and forecasting U.S. mortality. *Journal of the American Statistical Association, 87*(419), 659-671.

Leslie, W., & Hankey, C. (2015). Aging, nutritional status and health. *Healthcare (Basel, Switzerland)*, 3(3), 648-658.

Lopez-Arana, S., Avendano, M., van Lenthe, F. J., & Burdorf, A. (2014). Trends in overweight among women differ by occupational class: Results from 33 low- and middle-income countries in the period 1992-2009. *International Journal of Obesity (2005), 38*(1), 97-105.

Majer, I. M., Mackenbach, J. P., & van Baal, P. H. (2013). Time trends and forecasts of body mass index from repeated cross-sectional data: A different approach. *Statistics in Medicine, 32*(9), 1561-1571.

Mamun, A. A., & Finlay, J. E. (2014). Shifting of undernutrition to overnutrition and its determinants among women of reproductive ages in the 36 low to medium income countries. *Obesity Research & Clinical Practice, 9*(1), 75-86.

Martorell, R., Khan, L. K., Hughes, M. L., & Grummer-Strawn, L. M. (2000). Obesity in women from developing countries. *European Journal of Clinical Nutrition, 54*(3), 247-252.

McLaren, L. (2007). Socioeconomic status and obesity. *Epidemiologic Reviews, 29*, 29-48.

McPherson, K., Marsh, T., & Brown, M. (2007). *Foresight. tackling obesities: Future choices - modelling future trends in obesity & their impact on health*. London, UK: Government Office for Science.

Mendez, M. A., Monteiro, C. A., & Popkin, B. M. (2005). Overweight exceeds underweight among women in most developing countries. *The American Journal of Clinical Nutrition, 81*(3), 714-721.

Meydani, M. (2001). Nutrition interventions in aging and age-associated disease. *Annals of the New York Academy of Sciences, 928*, 226-235.

Mills, T. C. (2009). Forecasting obesity trends in England. *Journal of the Royal Statistical Society Series A, 172*(1), 107-117.

Miranda, J. J., Herrera, V. M., Chirinos, J. A., Gomez, L. F., Perel, P., Pichardo, R. et al. (2013). Major cardiovascular risk factors in Latin America: A comparison with the United States. The Latin American consortium of studies in obesity (LASO). *PloS One, 8*(1), e54056.

Monteiro, C. A., Conde, W. L., Lu, B., & Popkin, B. M. (2004a). Obesity and inequities in health in the developing world. *International Journal of Obesity and Related Metabolic Disorders, 28*(9), 1181-1186.

Monteiro, C. A., Conde, W. L., & Popkin, B. M. (2007). Income-specific trends in obesity in Brazil: 1975-2003. *American Journal of Public Health, 97*(10), 1808-1812.

Monteiro, C. A., Conde, W. L., & Popkin, B. M. (2001). Independent effects of income and education on the risk of obesity in the Brazilian adult population. *The Journal of Nutrition, 131*(3), 881S-886S.

Monteiro, C. A., Moura, E. C., Conde, W. L., & Popkin, B. M. (2004b). Socioeconomic status and obesity in adult populations of developing countries: A review. *Bulletin of the World Health Organization, 82*(12), 940-946.

Nandi, A., Sweet, E., Kawachi, I., Heymann, J., & Galea, S. (2014). Associations between macrolevel economic factors and weight distributions in low- and middle-income countries: A multilevel analysis of 200,000 adults in 40 countries. *American Journal of Public Health, 104*(2), e162-71.

NCD Risk Factor Collaboration. (n.d.). *Adiposity, global data.* Retrieved 4/4, 2106 from http://www.ncdrisc.org/d-adiposity.html

NCD Risk Factor Collaboration (NCD-RisC). (2016). Trends in adult body-mass index in 200 countries from 1975 to 2014: A pooled analysis of 1698 population-based measurement studies with 19.2 million participants. *Lancet (London, England), 387*(10026), 1377-1396.

Neuman, M., Finlay, J. E., Davey Smith, G., & Subramanian, S. V. (2011). The poor stay thinner: Stable socioeconomic gradients in BMI among women in lower- and middle-income countries. *The American Journal of Clinical Nutrition, 94*(5), 1348-1357.

Neuman, M., Kawachi, I., Gortmaker, S., & Subramanian, S. (2014). National economic development and disparities in body mass index: A cross-sectional study of data from 38 countries. *PloS One, 9*(6), e99327.

Neuman, M., Kawachi, I., Gortmaker, S., & Subramanian, S. V. (2013). Urban-rural differences in BMI in low- and middle-income countries: The role of socioeconomic status. *The American Journal of Clinical Nutrition, 97*(2), 428-436.

Ng, M., Fleming, T., Robinson, M., Thomson, B., Graetz, N., Margono, C. et al. (2014). Global, regional, and national prevalence of overweight and obesity in children and adults during 1980-2013: A systematic analysis for the global burden of disease study 2013. *Lancet (London, England), 384*(9945), 766-781.

O'Donnell, M. J., Chin, S. L., Rangarajan, S., Xavier, D., Liu, L., Zhang, H. et al. (2016). Global and regional effects of potentially modifiable risk factors associated with acute stroke in 32 countries (INTERSTROKE): A case-control study. *Lancet (London, England), 388*(10046), 761-775.

O'Donnell, M. J., Xavier, D., Liu, L., Zhang, H., Chin, S. L., Rao-Melacini, P. et al. (2010). Risk factors for ischaemic and intracerebral haemorrhagic stroke in 22 countries (the INTERSTROKE study): A case-control study. *Lancet, 376*(9735), 112-123.

Olaiz, G., Rojas, R., Barquera, S., Shamah, T., Aguilar, C., Cravioto, P. et al. (2003). *Encuesta nacional de salud 2000. tomo 2. la salud de los adultos*. Cuernavaca, México: Instituto Nacional de Salud Pública (INSP).

Olaiz-Fernández, G., Rivera-Dommarco, J., Shamah-Levy, T., Rojas, R., Villalpando-Hernández, S., Hernández-Avila, M. et al. (2006). *Encuesta nacional de salud y nutrición 2006*. Cuernavaca, Mexico: Instituto Nacional de Salud Pública (INSP).

Ouyang, Y., Wang, H., Su, C., Du, W., Wang, Z., & Zhang, B. (2015). Why is there gender disparity in the body mass index trends among adults in the 1997-2011 china health and nutrition surveys? *Asia Pacific Journal of Clinical Nutrition, 24*(4), 692-700.

Pampel, F. C., Denney, J. T., & Krueger, P. M. (2012). Obesity, SES, and economic development: A test of the reversal hypothesis. *Social Science & Medicine, 74*(7), 1073-1081.

Penman, A. D., & Johnson, W. D. (2006). The changing shape of the body mass index distribution curve in the population: Implications for public health policy to reduce the prevalence of adult obesity. *Preventing Chronic Disease, 3*(3), A74.

Perez-Escamilla, R., Villalpando, S., Shamah-Levy, T., & Mendez-Gomez Humaran, I. (2014). Household food insecurity, diabetes and hypertension among Mexican adults: Results from ENSANUT 2012. *Salud Publica de Mexico, 56 Suppl 1*, s62-70.

Perez Ferrer, C., McMunn, A., Rivera Dommarco, J. A., & Brunner, E. J. (2014). Educational inequalities in obesity among Mexican women: Time-trends from 1988 to 2012. *PloS One, 9*(3), e90195.

Philipson, T. J., & Posner, R. A. (2003). The long-run growth in obesity as a function of technological change. *Perspectives in Biology and Medicine, 46*(3 Suppl), S87-107.

Pischon, T., Boeing, H., Hoffmann, K., Bergmann, M., Schulze, M. B., Overvad, K. et al. (2008). General and abdominal adiposity and risk of death in Europe. *The New England Journal of Medicine, 359*(20), 2105-2120.

Popkin, B. M. (1994). The nutrition transition in low-income countries: An emerging crisis. *Nutrition Reviews, 52*(9), 285-298.

Popkin, B. M., Adair, L. S., & Ng, S. W. (2012). Global nutrition transition and the pandemic of obesity in developing countries. *Nutrition Reviews, 70*(1), 3-21.

Popkin, B. M., & Gordon-Larsen, P. (2004). The nutrition transition: Worldwide obesity dynamics and their determinants. *International Journal of Obesity and Related Metabolic Disorders, 28 Suppl 3*, S2-9.

Poterico, J. A., Stanojevic, S., Ruiz-Grosso, P., Bernabe-Ortiz, A., & Miranda, J. J. (2012). The association between socioeconomic status and obesity in Peruvian women. *Obesity (Silver Spring, Md.), 20*(11), 2283-2289.

Preston, S. H., Heuveline, P., & Guillot, M. (2001). *Demography: Measuring and modeling population processes*. Oxford, UK: Blackwell Publishers.

Profamilia, & ICF Macro. (2011b). *Encuesta nacional de demografía y salud (ENDS) 2010 [datasets]*. Bogotá, Colombia: Profamilia, and Calverton, USA: ICF Macro [producer]. Rockville, USA: ICF International [distributor]. https://dhsprogram.com/data.

Profamilia, & ICF Macro. (2011a). *Encuesta nacional de demografía y salud, ENDS 2010*. Bogotá, Colombia: Profamilia, and Calverton, USA: ICF Macro.

Profamilia, & Macro International. (2005b). *Encuesta nacional de demografía y salud (ENDS) 2005 [datasets]*. Bogotá, Colombia: Profamilia, and Calverton, USA: Macro International [producer]. Rockville, USA: ICF International [distributor]. https://dhsprogram.com/data.

Profamilia, & Macro International. (2005a). *Salud sexual y reproductive. resultados: Encuesta nacional de demografía y salud 2005*. Bogotá, Colombia: Profamilia, and Calverton, USA: Macro International.

Profamilia, & Macro International. (2000b). *Encuesta nacional de demografía y salud (ENDS) 2000 [datasets]*. Bogotá, Colombia: Profamilia, and Calverton, USA: Macro International [producer]. Rockville, USA: ICF International [distributor]. https://dhsprogram.com/data.

Profamilia, & Macro International. (2000a). *Salud sexual y reproductiva. resultados: Encuesta nacional de demografía y salud 2000*. Bogotá, Colombia: Profamilia, and Calverton, USA: Macro International.

Profamilia, & Macro International. (1995b). *Encuesta nacional de demografía y salud (ENDS) 1995 [datasets]*. Bogotá, Colombia: Profamilia, and Calverton, USA: Macro International [producer]. Rockville, USA: ICF International [distributor]. https://dhsprogram.com/data.

Profamilia, & Macro International. (1995a). *Encuesta nacional de demografía y salud 1995*. Bogotá, Colombia: Profamilia, and Calverton, USA: Macro International.

Quezada, A. D., & Lozada-Tequeanes, A. L. (2015). Time trends and sex differences in associations between socioeconomic status indicators and overweight-obesity in Mexico (2006-2012). *BMC Public Health, 15*, 1244-015-2608-2.

Quispe, R., Benziger, C. P., Bazo-Alvarez, J. C., Howe, L. D., Checkley, W., Gilman, R. H. et al. (2016). The relationship between socioeconomic status and CV risk factors: The CRONICAS cohort study of Peruvian adults. *Global Heart, 11*(1), 121-130.e2.

Rao, J. N. K., Wu, C. F. J., & Yue, K. (1992). Some recent work on resampling methods for complex surveys. *Survey Methodology, 18*(2), 209-217.

Razak, F., Corsi, D. J., & Subramanian, S. V. (2013). Change in the body mass index distribution for women: Analysis of surveys from 37 low- and middle-income countries. *PLoS Medicine, 10*(1), e1001367.

Rigby, R. A., & Stasinopoulos, D. M. (2005). Generalized additive models for location, scale and shape. *Journal of the Royal Statistical Society Series C Applied Statistics., 54*(3), 507-554.

Rigby, R. A., & Stasinopoulos, D. M. (2004). Smooth centile curves for skew and kurtotic data modelled using the box-cox power exponential distribution. *Statistics in Medicine, 23*(19), 3053-3076.

Ruhm, C. J. (2007). Current and future prevalence of obesity and severe obesity in the united states. *Forum for Health Economic & Policy, 10*(2), 1-26.

Rutstein, S. O., & Way, A. (2014). *The Peru continuous DHS experience. DHS occasional papers no. 8*. Rockville, USA: ICF International.

Rutstein, S., & Staveteig, S. (2014). *Making the demographic and health surveys wealth index comparable. DHS methodological reports no. 9.* Rockville, Maryland, USA: ICF International.

Rutstein, S. O. (2008). *The DHS wealth index: Approaches for rural and urban areas. DHS working papers. no. 60*. Calverton, Maryland, USA: Macro International.

Rutstein, S. O., & Kiersten, J. (2004). *The DHS wealth index. DHS comparative reports no. 6*. Calverton, Maryland, USA: ORC Macro.

Schargrodsky, H., Hernandez-Hernandez, R., Champagne, B. M., Silva, H., Vinueza, R., Silva Aycaguer, L. C. et al. (2008). CARMELA: Assessment of cardiovascular risk in seven Latin American cities. *The American Journal of Medicine, 121*(1), 58-65.

Schneider, H. J., Friedrich, N., Klotsche, J., Pieper, L., Nauck, M., John, U. et al. (2010). The predictive value of different measures of obesity for incident cardiovascular events and mortality. *The Journal of Clinical Endocrinology and Metabolism, 95*(4), 1777-1785.

Sepulveda-Amor, J., Angel Lezana, M., Tapia-Conyer, R., Luis Valdespino, J., Madrigal, H., & Kumate, J. (1990). Nutritional status of pre-school children and women in Mexico: Results of a probabilistic national survey. *Gaceta Medica de Mexico, 126*(3), 207-24; discussion 224-6.

Shamah-Levy, T., Villalpando-Hernández, S., & Rivera-Dommarco, J. A. (2007). *Resultados de nutrición de la ENSANUT 2006*. Cuernavaca, Mexico: Instituto Nacional de Salud Pública (INSP).

Sobal, J., & Stunkard, A. J. (1989). Socioeconomic status and obesity: A review of the literature. *Psychological Bulletin, 105*(2), 260-275.

Sperrin, M., Marshall, A. D., Higgins, V., Buchan, I. E., & Renehan, A. G. (2014). Slowing down of adult body mass index trend increases in England: A latent class analysis of cross-sectional surveys (1992-2010). *International Journal of Obesity, 38*(6), 818-824.

Stevens, G. A., Singh, G. M., Lu, Y., Danaei, G., Lin, J. K., Finucane, M. M. et al. (2012). National, regional, and global trends in adult overweight and obesity prevalences. *Population Health Metrics, 10*(1), 22-7954-10-22.

Subramanian, S. V., Perkins, J. M., Ozaltin, E., & Davey Smith, G. (2011). Weight of nations: A socioeconomic analysis of women in low- to middle-income countries. *The American Journal of Clinical Nutrition, 93*(2), 413-421.

Swinburn, B. A., Sacks, G., Hall, K. D., McPherson, K., Finegood, D. T., Moodie, M. L. et al. (2011). The global obesity pandemic: Shaped by global drivers and local environments. *Lancet (London, England), 378*(9793), 804-814.

The DHS Program, ICF International. (n.d.). *Wealth index construction.* Retrieved 11/29, 2015 from http://www.dhsprogram.com/topics/wealth-index/Wealth-Index-Construction.cfm

Tucker, K. L., & Buranapin, S. (2001). Nutrition and aging in developing countries. *The Journal of Nutrition, 131*(9), 2417S-23S.

United Nations (UN). (2015b). *World population prospects: The 2015 revision. File POP/15-2: Annual female population by five-year age group, major area, region and country, 1950-2100 (thousands). Estimates, 1950 - 2015 [dataset]*. New York, USA: UN.

United Nations (UN). (2015a). *World population prospects: The 2015 revision. File POP/15-2: Annual male population by five-year age group, major area, region and country, 1950-2100 (thousands). Estimates, 1950 - 2015 [dataset]*. New York, USA: UN.

Valdespino, J. L., Olaiz, G., López-Barajas, M. P., Mendoza, L., Palma, O., Velázquez, O. et al. *Encuesta nacional de salud 2000. 2003. tomo I. vivienda, población y utilización de servicios de salud*. Cuernavaca, México: Instituto Nacional de Salud Pública (INSP).

Van Hook, J., Altman, C. E., & Balistreri, K. S. (2013). Global patterns in overweight among children and mothers in less developed countries. *Public Health Nutrition, 16*(4), 573-581.

Wang, Y., Beydoun, M. A., Liang, L., Caballero, B., & Kumanyika, S. K. (2008). Will all americans become overweight or obese? estimating the progression and cost of the US obesity epidemic. *Obesity (Silver Spring, Md.), 16*(10), 2323-2330.

Wang, Y. C., Colditz, G. A., & Kuntz, K. M. (2007). Forecasting the obesity epidemic in the aging U.S. population. *Obesity (Silver Spring, Md.), 15*(11), 2855-2865.

Wang, Y. C., McPherson, K., Marsh, T., Gortmaker, S. L., & Brown, M. (2011). Health and economic burden of the projected obesity trends in the USA and the UK. *Lancet (London, England), 378*(9793), 815-825.

Webber, L., Kilpi, F., Marsh, T., Rtveladze, K., Brown, M., & McPherson, K. (2012). High rates of obesity and non-communicable diseases predicted across Latin America. *PloS One, 7*(8), e39589.

World Bank. (2016b). *Data, country classification, A short story*. Retrieved 6/18, 2016 from http://go.worldbank.org/U9BK7IA1J0

World Bank. (2016a). *Data, GNI per capita, atlas method (current US$).* Retrieved 6/18, 2016 from http://data.worldbank.org/indicator/NY.GNP.PCAP.CD

World Health Organization (WHO). (2011). *Global status report on noncommunicable diseases 2010*. Geneva, Switzerland: WHO.

World Health Organization (WHO). (2006). *WHO child growth standards: Length/height-for-age, weight-for-age, weight-for-length, weight-for-height and body mass index-for-age: Methods and development*. Geneva, Switzerland: WHO.

World Health Organization (WHO). (2000). *Obesity: Preventing and managing the global epidemic: Report of a WHO consultation. (WHO technical report series 894)*. Geneva, Switzerland: WHO.

Zaninotto, P., Head, J., Stamatakis, E., Wardle, H., & Mindell, J. (2009). Trends in obesity among adults in England from 1993 to 2004 by age and social class and projections of prevalence to 2012. *Journal of Epidemiology and Community Health, 63*(2), 140-146.

Curriculum vitae

Goro YAMADA

Born in Funabashi, Chiba, Japan, on March 11, 1970.

Education

Doctor of Philosophy (PhD) Aug 2011 – Oct 2016

Department of International Health, Global Disease Epidemiology and Control Program

Master of Health Science (MHS)

Department of Biostatistics

Johns Hopkins Bloomberg School of Public Health – Baltimore, Maryland

Master of Health Sciences (MHS) Aug 1997 – May 1999

Department of International Health, Johns Hopkins School of Hygiene and Public Health – Baltimore, Maryland

Master of Medical Sciences (MMS) Apr 1992 – Mar 1994

Institute of Basic Medical Sciences, University of Tsukuba – Tsukuba, Ibaragi, Japan

Bachelor of Science (BSc) Apr 1988 – Mar 1992

College of Biological Sciences, University of Tsukuba – Tsukuba, Ibaragi, Japan

Professional experience

Project Formulation Advisor Feb 2010 – Mar 2011

Bolivia Country Office, Japan International Cooperation Agency (JICA) – La Paz, Bolivia

Chief Advisor / Expert in Infectious Disease Control Nov 2007 – Jan 2010

Amhara Regional Infectious Disease Surveillance Project, JICA – Bahir Dar, Ethiopia

Technical Officer Mar 2006 – Sep 2007

Health Metrics Network (HMN), Headquarters, World Health Organization (WHO) – Geneva, Switzerland

Technical Officer Sep 2005 – Mar 2006

Equity in Health Team, Headquarters, WHO – Geneva, Switzerland

Project Formulation Advisor Feb 2005 – May 2005

Honduras Country Office, JICA – Tegucigalpa, Honduras Mar 2004 – Jul 2004

Expert in Epidemiology Dec 2001 – Dec 2003

Project for Strengthening Regional Health Network of Santa Cruz Department, Bolivia, JICA – Santa Cruz, Bolivia

Associate Expert in Public Health Sep 2000 – Dec 2001

Headquarters, Medical Cooperation Department, JICA – Tokyo, Japan

Research Associate / Quality Assurance Officer Jun 1999 – Aug 2000

Intern Sep 1998 – Mar 1999

ZVITAMBO Project, Department of International Health, JHSPH – Harare, Zimbabwe

Science Teacher, Atebubu Secondary School Apr 1994 – Sep 1996

Japan Overseas Cooperation Volunteers (JOCV), JICA – Atebubu, Brong-Ahafo, Ghana

Honors and awards

Recipient May 2015

The Nancy Stephens Student Support Fund, Johns Hopkins Bloomberg School of Public Health

Inductee May 1999

Alpha Chapter, Delta Omega Society, Johns Hopkins School of Hygiene and Public Health

Selected publications and reports

Melby MK, **Yamada G**, Surkan PJ. Inadequate gestational weight gain increases risk of small-for-gestational-age term birth in girls in Japan: A population-based cohort study. *Am J Hum Biol*. 2016. 28(5), 714-720.

Gupta S, **Yamada G**, Mpembeni R, Frumence G, Callaghan J, Stevenson R, Brandes N, Baqui A. Factors Associated with Four or More Antenatal Care Visits and Its Decline among Pregnant Women in Tanzania between 1999 and 2010. *PLoS ONE.* 2014; 9(7): e101893.

Baseline Survey Report, Amhara Regional Infectious Disease Surveillance (AmRids) Project. Amhara Regional Health Bureau / Japan International Cooperation Agency (JICA), 2008.

Using the HMN framework to assess country health information system, interim report July 2007 (report to the 8th Health Metrics Network (HMN) Board Meeting). HMN, WHO, 2007.

Status report: country HIS assessment activities, 2007 (report to the Technical Advisory Group Meeting, Health Metrics Network). HMN, WHO, 2007.

Approaches for systematic planning of development project <anti-HIV/AIDS measures>, Institute for International Cooperation, Japan International Cooperation Agency (JICA), 2002.

Yamada G. Master of Health Science Essay: Quality assessment of lymph node palpation and anthropometry in the ZVITAMBO Project, Harare, Zimbabwe. School of Hygiene and Public Health, Johns Hopkins University, 1999.

Yamada G, Hama H, Kasuya Y, Masaki T, and Goto K. Possible sources of endothelin-1 in damaged rat brain. *J Cardiovasc Pharmacol.* 1995; 26, Supple. 3: S486-S490.

Hama H, Kasuya Y, Sakurai T, **Yamada G**, Suzuki N, Masaki T, and Goto K. Role of endothelin-1 in astrocyte responses after acute brain damage. *J Neurosci Res.* 1997; 47(6): 590-602.

Yasuda N, Kasuya Y, **Yamada G**, Hama H, Masaki T, and Goto K. Loss of contractile activity of endothelin-1 induced by electrical field stimulation-generated free radicals. *Br J Pharmacol.* 1994; 113(1): 21-28.

Shiba R, Sakurai T, **Yamada G**, Morimoto H, Saito A, Masaki T, and Goto K. Cloning and expression of rat preproendothelin-3 cDNA. *Biocem Biophys Res Commun.* 1992; 186(1): 588-594.

Selected presentations

The following will be presented at the American Public Health Association (APHA) Annual Meeting and Expo, Denver, 2016.

Yamada G, Jones-Smith J, Castillo-Salgado C, Moulton LH. Differences in magnitude and rate of change in BMI distributions by geographic and socioeconomic factors: Mexico, Colombia, and Peru, 2005-2010.

Yamada G, Jones-Smith J, Castillo-Salgado C, Moulton LH. Obesity projection by modeling BMI distribution: Using national health survey data from Mexico, Colombia, and Peru, 1988-2013.

Melby MK, **Yamada G**, Surkan PJ. Can the effects of inadequate gestational weight gain on small-for-gestational age risk be buffered by sociocultural factors?