UNIVERSITY OF LEEDS

Microeconomic development: An investigation on labour markets and middle classes

Jose Martin Lima Velazquez

Submitted in accordance with the requirements for the degree of

Doctor of Philosophy

The University of Leeds

Leeds University Business School
Department of Economics

January 2019

The candidate confirms that the work submitted is his/her own, except where work which has formed part of jointly-authored publications has been included. The contribution of the candidate and the other authors to this work has been explicitly indicated below. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others.

An earlier version of chapter II, with fewer countries, indices and without the analysis of RBL areas, has been presented at the IARIW general conference 2016, held in Dresden (21-27 August), and went through refereeing process.

The conference paper is available online in the conference proceedings as: “Financial crises, the Great Recession and the middle class in the Americas, an empirical analysis” by Jose Martin Lima Velazquez and Gaston Yalonetzky, 34th IARIW General Conference.

Work attributable to me: Application of statistical techniques in datasets

Work attributable to my co-author: Description of methodologies

A different version of chapter IV, without concentring just on the middle class, has been presented at the Asian Meeting of the Econometric Society 2018, held in Seoul (21-23 July), and went through refereeing process.

The conference paper is available online in the conference proceedings as: “Alternative policy-effect estimations under quasi-experimental designs: An application in Mexico” by Kauski Chaudhuri, Martin Lima and Gaston Yalonetzky. AMES 2018.

Work attributable to me: creation and development of synthetic panel estimations, data analysis.

Work attributable to my co-authors: guidance, result interpretation.

Acknowledgements

I would like to express my deep gratitude to my academic supervisors, Dr Gaston Yalonetzky and Dr Kausik Chaudhuri, for all their unconditional support, wise guidance, friendship and patience during the last four years. Their enthusiasm for the academic research, passion for teaching and wisdom is an inspiring driver to become a better human being.

To my family, Maria de los Angeles Velazquez, Dr Martin Lima, and Dr Sylvia Strawa, for all their support, comprehension and love during this period. I would also like to thank to Maria Eugenia Velazquez, Carlos and Cecilia Doblado for all their encouragement.

To my friends and colleagues for all their support, encouragement, guidance and positive thoughts: Dr Maria Gavris, Dr Daniel Bonilla, Josias Franco, Hector Olivas, Dr Bianca Orsi, Roland Marposa, Wendy Sanchez, Joshua Cave, Andree Pedersen and Enrique Minor.

And special thanks to all the administrative staff at the University, particularly to David Courtney, Matthew Armstrong, Yawen Ho, and Hellen Graves.

This research was fully sponsored by the Mexican National Council for Sciences and Technology (CONACYT). I would like to thank Georgina Mendez Reyes, Maria Dolores Sanchez and Samuel Manterola for all their support at CONACYT.

Finally, I would like to thank and acknowledge all the important people who have a huge influence in my life and had passed in the last years while I was doing my studies: Dr Eduardo Rodriguez-Oreggia, Georgina Mendez Reyes, Dr Nelly Aguilera Aburto, Maria de la Soledad Velazquez and Rogelio Zarate Rugerio.

Abstract

The relevance of the middle class in social sciences is undoubted; understanding the middle class is fundamental to understand Latin America. Despite the potential importance of the middle class, there is no consensus either on its definition or its measurement. This thesis investigates on the definition of the middle class, in order to provide a more comprehensive understanding of this social group in Mexico, from an income (unidimensional) perspective in Latin American to a multidimensional definition for Mexico. Moreover, this work analyses the Mexican middle-class women's labour conditions when facing no-fault divorce, and explores the costs on informal loans middle-class individuals face when financially excluded. Most middle-class income definitions (relative, absolute and a proposed bipolarisation analysis) suggest an expansion of the middle class in Brazil, Chile, Mexico, Paraguay, Peru, and Uruguay between 2000 and 2015. Income definitions, nevertheless, do not reflect the multidimensional nature of the middle class. This work proposes a multidimensional definition of the middle class based on counting the middle-class characteristics households have (counting approach), finding that an income definition of $10 and $50 USD per day has a good fit in identifying the proposed multidimensional definition, compared to other income definitions (34.0% identification overlap). Based on a quasi-experimental technique, middle-class women exposed to no-fault divorce participate 3.5 percent points more on the labour market, reduce their participation in the informal sector by 3.1 percent points, and worked on average one hour more than non-exposed middle-class women. Finally, based on a propensity score matching estimation, middle-class individuals in urban areas excluded from formal financial services pay 7.3 percent points more on interest rates for informal loans than non-excluded individuals, despite this loans where granted by relatives, friends and co-workers of the borrower.

Table of Contents

List of Tables

List of Figures

Chapter I
Introduction

The middle class is a subject of debate all over the world: in the media (Crompton, 2008: p 15), in politics (Kenny, 2011), and in everyday talks (Boyle, 2013: p 9). In Latin America, the values and aspirations of the middle class are even portrayed in the soap operas (Ferreira et al., 2013; Chong and La Ferrara, 2009). Robert Solow noted “there is no shortage of talk about the middle class” (Estache and Leipziger, 2009: p 1).

The ubiquity of this discussion is not a coincidence. The relevance of the middle class in social sciences is undoubtedly (Chakravarty, 2015: p 1; Boyle, 2013: p 39; Ferreira et al., 2013). Since ancient Greece, Aristotle, echoing Phocylides, highlighted the importance of an ample middle class to build a healthy society by balancing the extremes: the poor and the rich (Aristotle, 2017: p 99).

Modern research finds similar conclusions to Aristotle’s. The middle class works as a moderator between the rich and the poor (Thurow, 1984). A solid middle class reduces the risk of political instability, social tensions or conflict (Chakravarty, 2015: p1). A polarised society is prone to social tension, to the possibility of revolution or revolt, and generalised social unrest (Esteban and Ray, 1994).

Despite the potential importance of the middle class, there is no consensus either on its definition or its measurement. The interest for analysing the middle class arose from Sociology where, between the 1970s and 1980s, a debate took place between two perspectives of analysis: an interpretative (hermeneutic), and a positivistic (factual, binary) perspective (Crompton, 2008: p 17).

Economists got interested in the positivistic perspective in the early 1980s, as a way to measure the shrinkage or expansion of the middle class in the USA following economic crises (Thurow, 1984; Levy, 1987). After this period, the literature compiled an extensive catalogue of middle-class definitions, but nothing universal (Reeves et al., 2018). The fact that different measurement approaches could lead to different results (Barcena and Canto, 2018), despite analysing the same wellbeing domain (e.g. income), fuelled the discussion around the middle-class.

Extreme inequalities historically observed in Latin America led sociologists to believe in the absence of a middle class in the region, and in a binary social structure of top rich and bottom poor instead (Witheford, 2013). However, Daza and Cortes (2013) argue for the existence of a Latin American middle class, but the search for an appropriate definition is essential.

The economic analysis of the middle class has focused on a unidimensional, single wellbeing domain. This approach helps in the analysis of international and intertemporal comparisons (Birdsall, 2010; Kenny, 2011). However, the middle class has characteristics which go beyond a single dimension (Parker, 2013); being in the “middle” of a single wellbeing domain’s distribution does not capture all other characteristics which could define (and affect) the middle class (Castellani et al., 2015). A more comprehensive analysis of people's wellbeing could, arguably should, be multidimensional in nature (Alkire and Foster, 2011).

Labour market participation, and job characteristics, are a relevant dimension of the middle class (Parker, 2013; Crompton, 2008: p 16; Lopez, 2013). Another important dimension of the middle class is access to credit, as a source of resources to afford capital assets and properties (Organisation for Economic Cooperation and Development, OECD, 2010), and to smooth consumption over time and protecting households against adverse shocks (Zeller et al., 1997: p 12).

Considering all these arguments, this thesis attempts to provide a more comprehensive understanding of the middle class in Mexico, from a unidimensional income perspective within the Latin American context, to a multidimensional analysis of the Mexican middle class; as well as understanding middle-class woman labour conditions when facing the introduction of no-fault divorce, and exploring the real cost for informal loans faced by middle-class individuals financially excluded.

I.1 The analysis on the middle class

This thesis has four empirical chapters. Chapter II, on the middle class in Latin America, explores different unidimensional definitions of the middle class in a set of six Latin American countries (Brazil, Chile, Mexico, Paraguay, Peru, and Uruguay), considering the monthly household per capita income as a wellbeing domain. The middle-class definitions are based in relative (percentages of the median) and absolute (cut-offs) threshold settings, as well

as on bipolarisation analysis, which seeks to represent the middle class as the egalitarian opposite of bimodal distributions.

Chapter III, on the middle class in Mexico, develops a multidimensional definition of the middle class based on the counting approach (a method which ads or "counts" particular characteristics of a specific group), popular in the measurement of social exclusion and poverty, but novel in the analysis of the middle class. This definition is non-income based and considers five dimensions: education, access to health services, dwelling characteristics, household assets, and occupational characteristics. This chapter also presents a specific case of the counting approach based on multidimensional poverty measurements.

Chapter IV, “No-fault divorce and female labour market participation: evidence from the Mexican middle class”, assesses the impact of the no-fault divorce introduction in Mexico on middle-class women’s labour participation, as well as their participation in the informal labour sector, and their working hours, using a quasi-experimental estimation design.

Chapter V, on the real cost of financial exclusion among the Mexican middle class, investigates the differences in interest rates of informal loans between middle-class individuals financially excluded and non-excluded, as well as for the other social classes as a comparison (the deprived and the affluent).

I.2 Motivation

Latin America is the region with the most unequal income distribution worldwide (Parker, 2013); hence, the relevance of the middle-class analysis focused on the Latin American context (Daza and Cortes, 2013). The understanding of the Latin American middle class is fundamental to understanding contemporary society in the region (Whiteford, 2013).

The main motivation behind chapter II is to investigate the effects of, and differences between, alternative measurement approaches on the unidimensional definition and quantification of the middle class, by discussing different middle-class measurement approaches available in the literature, highlighting the controversy around fixing middle-class thresholds, and to contrast those approaches to proposed indices of bipolarization analysis. Moreover, this chapter also provides a contextual frame on the conditions of the middle class in Latin America to embed the Mexican case.

However, as discussed before, the middle class is rather a multidimensional phenomenon. The motivation of chapter III is to propose a non-income based middle-class definition which considers the multidimensional nature of wellbeing (Maasoumi and Yalonetzky, 2013; Alkire and Foster, 2011) and reflects the characteristics and identity of the middle class beyond the single economic domain (Parker, 2013; Wagley, 2013).

As mentioned before, labour market participation, and job characteristics, are a relevant dimension of the middle class (Parker, 2013; Crompton, 2008: p 16; Lopez, 2013). The motivation behind chapter IV is to provide evidence for Mexico of the effects of introducing no-fault divorce on middle-class women's labour participation, as well as in their labour characteristics. These findings could be transferred to the particular case of other developing countries when facing similar conditions.

Considering the relevance of credit as one of the main sources for household capital accumulation, protection against shocks, and consumption patterns of the middle class (OECD, 2010; Zeller et al., 1997: p 12), the motivation behind chapter V is to understand the financial exclusion in the middle class, as well as to test for differences in interest rates between middle-class individuals financially excluded and non-excluded, in order to assess the cost of financial exclusion.

I.3 Research questions

Three research questions are behind chapter II:

1. Is the middle class in Latin America shrinking or expanding within the period 2000 and 2015?

Following the debate in the USA after the economic crises between the 1970s and 1980s (Thurow, 1984; Levy, 1987), this question can be applied in the Latin American context after the world financial crisis of 2008. Most evidence suggests an expansion of the Latin American income middle class within this period (Ferreira et al., 2013; OECD, 2011). However, evidence could not be robust when applying different middle-class definitions and methods.

2. Are traditional unidimensional definitions of the middle class consistent in Latin America?

Based on the literature (Barcena and Canto, 2018; Levy, 1987), discrepancies in trends between different definitions are usually observed in developed

countries, suggesting that similar patterns could be found in the case of Latin American countries.

3. Are traditional unidimensional definitions consistent with estimations based on bipolarisation analysis?

The use of indices based on bipolarisation analysis to identify trends in the evolution of the middle class is not common among different countries in Latin America. Therefore, this analysis could help to understand the empirical consistency between traditional estimations of the middle class and bipolarisation analysis in the region.

The research question behind chapter III is:

1. Are unidimensional (income) and multidimensional (non-income) middle-class definitions consistent in the identification of the middle class?

Findings in Colombia by Castellani et al. (2015) demonstrate that both approaches do not identify the same individuals, leading to identification discrepancies. Same discrepancies are expected in this analysis for Mexico. However, a reasonable degree of overlap between definitions could help to identify a particular definition of income middle-class which reflects the multidimensionality of the phenomenon's characteristics. On the other hand, these discrepancies could be used to identify subgroups within the middle class (lower and upper middle class for instance).

Three research questions are behind chapter IV:

1. Is the no-fault divorce introduction affecting the middle-class women's labour market participation?

Following the Mexican National Human Rights Commission and the Supreme Court of Justice (CNDH, 2013; SCJN, 2012), the introduction of no-fault divorce would, in theory, enhance women's autonomy and middle-class female labour participation.

2. Is the no-fault divorce introduction affecting the middle-class women's participation in the informal sector[1]?

Following Parkman (1992), narrow definition of property rights and limited alimony (weak institutional protection of women) have a significant positive influence on women's labour participation in the USA after the introduction of

[1] Informal labour market lacks social security coverage following Mexico's Labour Law.

no-fault divorce. Considering the lack of a welfare state in Mexico, and its institutional limitations, a reduction in informality is expected in the aftermath of the introduction of no-fault divorce.

3. Is the introduction of no-fault divorce affecting middle-class women's working hours?

An increase in working hours is expected due to women's autonomy enhancement (CNDH, 2013), and as a protection against the absence of social security safety nets (Parkman, 1992).

Two research question are behind chapter V:

1. Are real interest rates for informal loans different between financially excluded and non-excluded middle-class individuals?

Considering that only loans between closely related individuals are taken into consideration, no differences in interest rates between excluded and non-excluded individuals should be expected, because borrowers and lenders are so close to each other that there should be no information asymmetries (Bloom et al., 2008; Lin et al., 2009; 2013).

2. Are real interest rates for informal loans different between financially excluded and non-excluded individuals from the deprived and affluent social classes?

Following the same reasoning as in the previous question, no differences between excluded and non-excluded individuals should be expected due to the minimisation of information asymmetries (Bloom et al., 2008; Lin et al., 2009; 2013).

I.4 Contributions

Chapter II contributes empirically by measuring different unidimensional middle-class definitions (relative and absolute), based on the per capita household income as wellbeing domain for six Latin American countries, and then contrasting these measurements to bipolarization definitions. A particular methodological contribution of this chapter is the identification of the areas below the Relative Bipolarization Lorenz curve proposed by Yalonetzky (2014), as indices of bipolarisation useful to infer changes in the middle class.

Chapter III contributes methodologically with a novel proposal to measure the middle-class based on the counting approach, popular for poverty identification (Castellani et al., 2005; Banerjeer and Duflo, 2008; Alkire et al.,

2015) but not used previously to identify the middle class. The effort includes operationalisation proposals involving how to choose a relevant dimension and how to treat them regarding their ability to distinguish the middle class from both the deprived and the affluent. Finally, the empirical implementation of these methods and its contrast to traditional unidimensional definitions is deemed a relevant empirical contribution to the literature.

Chapter IV analyses the impact of no-fault divorce on middle-class female labour market behaviour; a topic which has been previously studied in developed countries, but not yet in developing countries. Moreover, no previous studies have analysed these effects on women from the middle class, making the analysis in this chapter a pioneer in the field. A novel methodological contribution of this chapter is the proposal of synthetic panels in the difference-in-difference estimation.

Finally, chapter V contributes to the existing literature in several ways: first, there are no previous studies in the literature using information on both amounts borrowed and paid from informal loans, thus enabling to define real interest rate directly in this context. Second, this chapter considers the trust between individuals as a control for information screening, which is not common in the existing literature. And finally, no previous analysis has focused on the differences in real interest paid across social classes.

I.5 Thesis outline

This thesis is divided in six chapters. Chapter II explores unidimensional definitions of the middle class in Latin American. Chapter III develops a multidimensional definition of the middle class in Mexico. Chapter IV assesses the impact of the no-fault divorce introduction in Mexico on middle-class women's labour participation and characteristics. Chapter V estimates the real cost of financial exclusion among the Mexican middle class between individuals financially excluded and non-excluded. Finally, chapter VI presents the conclusions of this thesis, discussing the findings of all these chapters, and closes this work with a future agenda.

Chapter II
The middle class in Latin America[2]

II.1 Introduction

This chapter discusses different definitions of, and approaches to, the middle class in Latin America, and presents an empirical analysis of its evolution, based on a unidimensional approach in terms of the per capita household income as wellbeing domain, during the period 2000-2015, for six countries[3]: Brazil, Chile, Mexico, Paraguay, Peru, and Uruguay.

The main motivation behind this chapter is to investigate the effects of, and differences between, alternative measurement approaches on the unidimensional definition and quantification of the middle class. The main objectives of this chapter are to discuss the different middle-class measurement approaches available in the literature, highlighting the controversy around fixing middle-class thresholds, and contrasting those approaches to the bipolarisation analysis (which also seeks to represent the middle class as the egalitarian opposite of bimodal distributions).

Latin America is a region historically affected by social inequalities (the most unequal region in the world according to Parker, 2013). This situation leads Sociologists to believe Latin America has no middle class but only a binary social structure of top rich and bottom poor (Whiteford, 2013). Daza and Cortes (2013) argue the existence of a middle class in Latin America, but the search for an appropriate definition of the middle class in this region plays an essential role in the discussion. Hence the relevance of this work.

The research questions of this chapter are exploring potential discrepancies among traditional unidimensional definitions of the middle class in the Latin American region, the comparison of these definitions to the bipolarisation analysis, in particular to the methodology proposed in this chapter, and to test whether the middle class is shrinking or expanding in the region between 2000 and 2015.

---

2 An earlier version of this chapter, with fewer countries, indices and without the analysis of RBL areas, has been presented at the IARIW general conference 2016, held in Dresden (21-27 August), and went through refereeing process.

3 The selection of these countries as well as the time frame obey data availability. These restrictions will be discussed in section II.3.2.

The empirical analysis of this chapter confirms significant discrepancies between estimations based on different middle-class definitions and shows that these variations are explained by the relative arbitrariness involved in defining their thresholds. Then, the relevance of this analysis on the literature is the exploration of less arbitrary middle-class estimation approaches like bipolarisation analysis, which was introduced by Foster and Wolfson (1992, 2010) as a critique to the arbitrariness of fixed middle-class thresholds.

A particular methodological contribution of this chapter is the identification of the areas below the Relative Bipolarisation Lorenz curve proposed by Yalonetzky (2014). These areas can be used as indices of bipolarisation, so they are useful to infer changes in the middle class.

The analysis of the middle class in Latin America helps to contextualise the Mexican case in the region, and to embed further analysis of the Mexican middle class developed in chapters IV, and V of this thesis; while chapter III complements the analysis presented in this chapter by elaborating a detailed study and providing empirical estimations based on the multidimensional identification of the middle class.

This chapter is divided in six sections. The following section builds on the definition of the middle class from a literature review. The third section explores different methods to estimate the middle class, focusing on the unidimensional case[4], for three definitions: relative measurements, absolute measurements, and bipolarisation analysis. The fourth section gives a brief and general description of economic conditions in the sample of countries. The fifth section shows the empirical analysis, describing the data used for these estimations and results. Finally, a sixth section concludes this chapter with a discussion on the findings.

II.2 Defining the middle class: a review of the literature

This section explores the definition of the middle class without attempting to build a new ontology on this concept, but just visiting the literature. Thus, this discussion could be considered as a theoretical foundation on which further methodological discussions are supported.

---

4 As mentioned before, further discussion on the multidimensional case is elaborated in chapter III.

II.2.1 The relevance of the middle class

The relevance of the middle class in societies is not a recent concern; it has a well-documented history on its influence in many aspects of social life (Savage, 2015: p 25). An early reference could be traced back to ancient Greece where Aristotle discussed the importance of the middle class for a balanced political system in order to prevent *hyperbolas* (opposing excesses). Moreover, he claimed that “cities can be well governed [when] the middle class is numerous and stronger” (Aristotle, 2017: 99).

Then, the relevance of the middle class is undoubtedly in many aspects of modern societies. Banerjee and Duflo (2008), after a comprehensive literature review, identify three key arguments on the relevance of the middle class:

1. Provides new and more patient entrepreneurs who will create labour opportunities.
2. Has a particular set of values which emphasise the accumulation of human capital and savings.
3. Provides costumers willing to pay a bit more for extra quality.

Regarding social conflict, the middle class works as a moderator between the rich and the poor (Thurow, 1984). Chakravarty (2015: p1) affirms that a solid middle class reduces the risk of political instability, social tensions or conflict. For Esteban and Ray (1994: p 819), social polarisation “is closely related to the generation of social tensions, to the possibilities of revolution and revolt, and to the existence of social unrest in general”. Hence the relevance of analysing social polarisation in societies.

Concerning governance, Loayza et al. (2012) demonstrated a relevant link between a mature middle class and its influence on good political practices by diminishing corruption. A report by the Organization of Economic Co-operation and Development (OECD) highlights the influence the middle class can have on the fiscal policy due to electoral preferences (OECD, 2010). For Birdsall (2010) and Graham and Pettinato (2001), the middle class could be considered as the backbone of both market economy and democracy in most advanced countries.

The middle class is also relevant in institutional terms. Acemoglu and Robinson (2009: p 256) suggested that the size of the middle class may explain why countries like Colombia or Costa Rica have a long history of uninterrupted democracy compared to their neighbours. They claim that a large middle class may act as a buffer against social conflict fed by potential demands of high level of redistribution by the poor, and by the temptation to face them with political repression by the rich.

The middle class is an important source of skilled workers for the labour market, playing an essential role in the development of industrial economies (Foster and Wolfson, 2010). Individuals from middle-class backgrounds are usually highly educated, and perform activities related to arts and design, as well as administrative, governmental and professional (Boyle, 2013: p 41).

The relationship between the middle class and the economic growth could be explained in two ways simultaneously: as an outcome and as an input of economic growth. This virtuous cycle could be extended to democracy and institutional development (Birdsall, 2010). Empirical evidence suggests that countries with larger middle class grow faster than countries with smaller middle class when ethnical diversity is not present (Easterly, 2001).

II.2.2 The middle class identity

Despite the well-documented relevance of this social group, there is no consensus about its definition in either sociological, historical or economic literature, leading to debate among scholars. Atkinson and Brandolini (2011) describe a kind of “penumbra” surrounding this concept. For Ferreira et al. (2012) the middle class is a slippery and multifaceted concept challenging to frame.

In his historical study, Marx (2015: p 56) identified the middle class during the Middle Ages as the independent craftsmen, the traders, and the free population owning their labour and means of production. Because means of production and consumption did not change over centuries, this definition could be consistently identifiable among different societies. This social group, however, transformed during the Industrial Revolution to the owners of the means of production, known as the bourgeois.

Social classes, nevertheless, could not only be defined by the ownership of means of production in modern societies. Conspicuous consumption describes a type of consumption made by the affluent in a society, the leisure class, which establishes a fashion or a mode, a “must be” of social acceptance, and could permeate to lower social groups, the impecunious classes, through aspirational drivers in a process called pecuniary emulation (Veblen, 2009: p 57).

Examples of pecuniary consumption and the middle class are Parker (2013), who found that white-collar workers in the 1920s Peru would demand an increase in salaries before decreasing their living standards. In the case of

mid-20th century Brazil, the middle class was closely related to the possession of goods which working classes would not need, deserve or expect (Owensby, 2013). Currid-Halkett (2017: p11) describes how brand logos in garments and cars could work as class determinants.

Pecuniary emulation, nevertheless, is not the only driver of consumption patterns. The process of mass production and technological progress observed after the Second World War allowed the popularisation and dissemination of new conspicuous goods in a process of “luxury democratisation" (Currid-Halkett, 2017: p 10).

Ethnic characteristics and the personal context could also define the middle-class membership. Wagley (2013) raises the relevance of ethnic backgrounds (European descendants, mixed-ethnical groups, local Native Americans), migration condition, and even areas of residence differentiated by social class in Latin America.

Broader definitions of the middle class also consider labour status[5] (employed, unemployed, self-employed), type of employment (skilled or unskilled), labour informality, level of education, access to financial services[6], household and household' asset ownership, pension coverage, and attitudes to entrepreneurship, among others (Banerjee and Duflo, 2008; OECD, 2010; Parker, 2013). The following chapters discuss some of these definitions in-depth.

Motivated by the economic crises experienced during the 1970’s, economists looked into the definition of the middle class with particular interest, a field traditionally reserved for Sociology (Crompton, 2008: p17), in order to see whether this social group was shrinking or expanding as an effect of the economic crisis at that time (Levy, 1987). Since then, economists have defined the middle class regarding wellbeing conditions, usually characterised as household income (positive), in opposition to sociological approaches focused on socio-cultural characteristics (hermeneutics).

A limitation to the economic approach to the middle class is that definitions based on the “middle mass”, a middle sector in terms of a wellbeing distribution, do not reflect the shared attributes and social relations

---

5 Chapter III develops on the labour dimension of the middle class.

6 Chapter V develops on the relationship of the Mexican middle class and credit access.

established within these groups (Wagley, 2013), nor identifies non-monetary vulnerabilities faced by the middle class (Castellani et al., 2015).

The central debate among economists concerns the approaches followed to define those characteristics which determine if individuals or households belong into this social class, and the establishment of boundaries differentiating the middle class from other social groups.

II.2.3 Consistency of the middle class over time

Fixing a single basket of goods, services or socio-economic characteristics to identify the middle class in a particular point in time faces the challenge of constant change over time due to changes in preferences, technological innovation or improvements in living conditions. Time permanence is the main difficulty in constructing multifaceted definitions of the middle class. To trace its changes over time, these definitions have to consider the dynamic nature of the middle class.

Kenny (2011: p 5) remarks on this problem: “it is hard to find a set of characteristics or values that are consistently and uniquely middle class across countries and time”. Parker (2013) highlights the relevance of middle-class definitions depending on understandings in their own time and place. Dayton-Johnson (2015) justifies the use of concrete economic definitions of the middle class (thresholds), in order to track its changes over time. However, this approach is mainly quantitative, and should be complementary to other social sciences’.

II.3 Measuring the middle class

Most economic literature discusses the measurement and identification of poverty due to its relevance in public policy and development. Therefore, full attention is paid to the definition of poverty thresholds, but not to the middle class as well as to the affluent group.

An early approach to the estimation of the middle class is found in Cole (1950), who argues that for statistical measurement of the middle class, the first step is to define the group’s characteristics. Then, estimations must be computed based on the available demographic data. However, although this method could fulfil all the objectivity of a positivistic measurement, results could vary among researchers due to the definitions they employ.

Lately, the measurement of the rich has also gained prominence with the work of Atkinson et al. (2016), who analysed the wealthy-elite in order to understand its influence in London's public life. Piketty also analyses the historical evolution of the top-rich as a way to understand the current concentration of wealth and to explain the structure of inequality in the world (Piketty, 2014: 254), and for the particular case of the USA (Piketty and Saez, 2003).

However, in the case of the middle-class identification, a challenge is presented when defining lower and upper thresholds to identify deprived and affluent population simultaneously. These challenges are one of the main motivations of this chapter: analysing the effects of setting poverty and affluence thresholds simultaneously, looking beyond traditional poverty estimations.

As mentioned in section II.2, the definition of the middle class could depend on political, social and cultural conditions, besides consumption patterns and fashion (including conspicuous consumption behaviours). Therefore, a time-invariant multidimensional definition of the middle class to allow comparisons between countries could be difficult to operationalise. In this sense, unidimensional income definitions, particularly relative to a certain point of a wellbeing distribution, enable the identification of the middle class by country and time without generalisations (Birdsall, 2010; Ferreira et al., 2013).

On the other hand, Castellani et al. (2015) provide some examples of multidimensional estimations, wherein the Colombian case the middle class is identified as the population not regarded as poor according to a multidimensional poverty condition, and income below 150% of the national income median. Edo et al. (2017) define the middle class in Argentina considering the distribution of multiple dimensions of welfare. Atkinson and Brandolini (2011) produce a definition of middle class for selected developed countries based on occupation characteristics.

A significant limitation of multidimensionality-based studies is the comparability among countries. In the case of Castellani et al. (2015), their definition of poverty binds exclusively to Colombia, with an *ad-hoc* household survey. The multivariate quantiles defined by Edo et al. (2017) are restricted to the Argentinian distribution and subject to changes in different contexts. Atkinson and Brandolini (2011) express concerns about the reliability of labour market status due to idiosyncratic national biases (e.g. differences in employment classifications between countries).

Taking these difficulties into consideration, a more common approach to define the middle class is relying on a single, comparable wellbeing domain. Income is a natural metric on a single dimension which approaches consumption; whereas expenditure represents individuals' consumption more accurately (Foster et al., 2013: p 4). Expenditure is also a more stable wellbeing indicator over time because it reacts slowly to changes in current income, drawing a smooth trend over time called "consumption inertia" (Deaton and Muellbauer, 1999: p 329).

Using a unidimensional wellbeing variable makes it easier to identify lower, middle and upper groups in a given distribution (Ferreira, et al., 2013: p 31). Foster and Wolfson (1992; 2010) identify, from previous empirical analysis of the middle class, four steps to identify this social group:

1. **Space selection**: selection of the welfare indicator, a wellbeing domain to work. A common choice is the income space, represented by salary or household incomes. It could also be defined as expenditure, or even as a people space in the case of population percentiles (Levy, 1987).
2. **Definition of the middle**: setting the appropriate definition of the middle group, based on its characteristics.
3. **Fixing the range**: establishing the thresholds or boundaries in which the middle class lies in terms of the wellbeing domain.
4. **Data aggregation**: aggregation of the middle-class information into an index which shows its magnitude in the domain of interest.

Following these steps, early economic definitions of the middle class fixed a range on the income distribution with lower and upper thresholds, differentiating the deprived of the affluent groups (Thurow, 1984; Blackburn and Bloom, 1985; Levy, 1987). The individuals lying within these boundaries are considered the middle class.

A key element to consider in this approach is the definition of these boundaries, which could be based on hard lines defined according to certain desirable characteristics (e.g. poverty lines), or according to characteristics of the populations' distribution on the wellbeing domain. The former approach is known as the absolute definition, and is preferred for international comparisons when using standardized units like USD purchasing power parity (PPP) (Ferreira, et al., 2013: p 31); whereas the latter is referred to as relative definitions and, as mentioned before, is preferred for comparisons over time (Birdsall, 2010).

All absolute and relative measurements, nevertheless, are subject to debate due to the arbitrariness of the threshold selection[7]. Moreover, they could lead to different results even when using the same wellbeing domain. Barcena and Canto (2018) document the discussion between Levy (1987) and Beach (1989) about the assumption of the former, and the scepticism of the later, on the vanishing of the middle class in the USA. Using the same data (US census and surveys) both demonstrate opposite conclusions just by using different definitions.

Foster and Wolfson (1992; 2010) raised awareness about the arbitrariness of the range fixing and suggested the use of bipolarisation approaches instead. Following the notion of convergence clubs (Quah, 1996), and the line of thought of Esteban and Ray (1994), the bipolarisation approach estimates the degree of distributional bimodality around the median, giving an intuitive approximation to the size of the middle class and its evolution. In this way, a bipolarised society would tend to concentrate on the extremes of the domain, whereas a society with larger middle class will concentrate around the centre.

The following subsections explore and discuss the relative, absolute and bipolarisation definitions of the middle class for the income[8] unidimensional[9] case. Many threshold-specific definitions are available in academic publications. However, the ones analysed in this chapter were chosen due to their relevance in the literature, as well as their adoption by international organisations like the World Bank or the OECD as official measures.

II.3.1 Relative definitions

Some of the earlier economic definitions of the middle class considered thresholds related to the median of a per capita distribution of a given wellbeing domain[10] (Ferreira, et al., 2013: p 31). Thurow (1984) defines the middle class as the population group with per-capita income between 75% and 125% of the income median (this definition is followed by Birdsall et al.,

---

7 The "surrounding penumbra" mentioned by Atkinson and Brandolini (2011).

8 Chapter III presents a detailed discussion on the differences between the use of income and expenditure as wellbeing proxy.

9 An exposition of the multidimensional case is developed in chapter III.

10 Mainly per-capita income.

2000); whereas Blackburn and Bloom (1985) give a higher threshold to the affluent, defining the middle class at 60% and 225% of the median.

Davis and Huston (1992) define the middle class as the population group with per capita income within 50% and 150% of the median. The OECD adopted this definition as its official international middle-class thresholds due to the widespread use of the 50% cut-off as a low income/poverty line in academic studies (OECD, 2010: p 59).

Another approach is based on people space definitions, measured as ranks or positions within the domains' distribution (Ferreira, et al., 2013: p 31), which could be income standards expressed as distributional percentiles (Foster et al., 2013: p 5). Levy (1987) defines the middle class as individuals with per capita income within the 20th and 80th percentiles (this definition is followed by Barro, 1999; and Easterly, 2001), whereas Partridge (1997) and Solimano (2008) define the middle class as the population between 40th and 60th percentiles, and between 20th and 90th percentiles, respectively. Table II.1 shows the thresholds for these relative definitions.

Table II.1 Thresholds for the relative definitions of the middle class

Relative definitions	Middle class range		
Based on percentage of the median	**Lower threshold**		**Upper threshold**
Thurow (1984), Birdsall et al (2000)	0.75 * y(p50)	y_i	1.25 * y(p50)
Blackburn and Bloom (1985)	0.60 * y(p50)	y_i	2.25 * y(p50)
Davis and Hudson (1992), OECD (2010)	0.50 * y(p50)	y_i	1.50 * y(p50)
Based on percentiles			
Levy (1987), Barro (1999), Easterly (2001)	y(p20)	y_i	y(p80)
Partridge (1997)	y(p40)	y_i	y(p60)
Solimano (2008)	y(p20)	y_i	y(p90)

Source: based on Ferreira, et al. (2013: p 31), Dayton-Johnson (2015), OECD (2010: p 59).
Note: p stands for the population percentile of the wellbeing domain distribution, p50 being the median.

II.3.2 Absolute definitions

As mentioned before, absolute definitions of the middle class are hard thresholds set according to certain desirable characteristics of this group. An example of these thresholds is poverty lines, defined to guarantee a minimum level of consumption. In the case of the middle class, however, an upper threshold is also set in order to differentiate the middle class from the affluent. This yields the challenge of setting two thresholds simultaneously.

As mentioned before, thresholds based on USD PPP are widely used for international comparisons. Milanovic and Yitzhaki (2002) set the middle-class

thresholds at $12 and $50 USD PPP per day. Banerjee and Duflo (2008) set these thresholds at $2 and $10 USD PPP per day. Ravallion (2010) sets the thresholds at $2 and $13 USD PPP per day[11].

Lopez-Calva and Ortiz-Juarez (2014) proposed a $10 and $50 USD PPP threshold. This definition, however, was tested for consistency through different countries in Latin America in order to estimate the likelihood of falling into poverty based on a set of household characteristics, such as asset ownership, which are associated with movements in or out of income poverty.

Their findings suggest that a 10% of probability of falling into income poverty is a good approximation of vulnerability for most countries in the region. The income level at this condition is close to $10 USD PPP per day. Therefore, the World Bank uses these thresholds for their middle-class estimations in Latin American. Table II.2 shows these thresholds for the absolute definitions.

Table II.2 Thresholds for the middle class relative definitions

Absolute definitions	**Middle class range**		
USD, PPP per day	**Lower threshold**		**Upper threshold**
Milanovic and Yitzhaki (2002)	$12	y_i	$50
Banerjee and Duflo (2008)	$2	y_i	$10
Ravallion (2010)	$2	y_i	$13
Lopez-Calva and Ortiz-Juarez (2011), World Bank (2013)	$10	y_i	$50

Source: based on Ferreira, et al. (2013: p 31), Dayton-Johnson (2015), OECD (2010: p 59).

II.3.3 Definitions based on bipolarisation analysis

Foster and Wolfson (1992) raised attention over the arbitrariness of the definition and measurement of the middle class in both relative and absolute approaches. This issue concerns the diversity of possible results obtained from a single distribution just from changing those thresholds. Their analysis focused on measuring changes in the middle class over time in the USA and Canada, using bipolarisation approaches instead of relative or absolute ones.

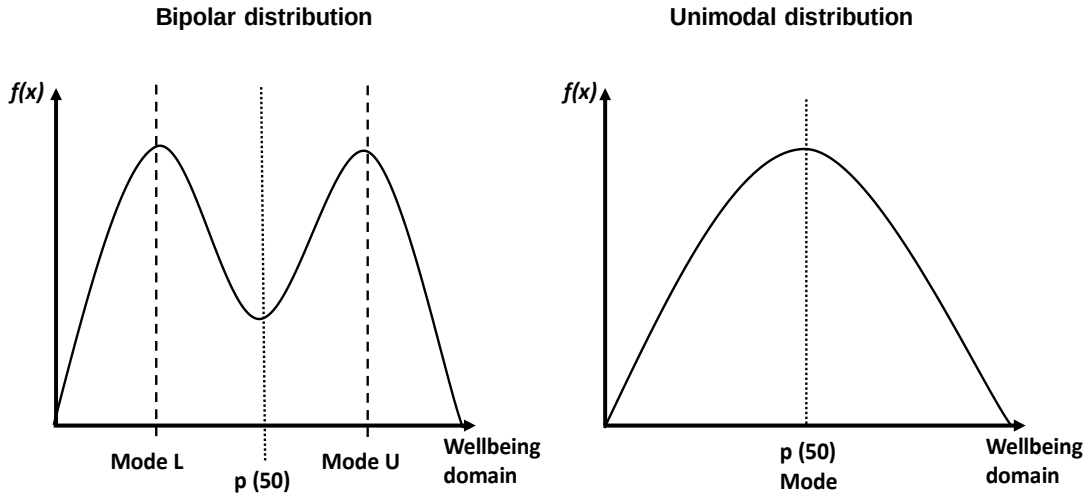
In his seminal paper, Quah (1996), using bipolarisation analysis, demonstrates that countries were clustering according to similar levels of development, a notion called convergence clubs, instead of converging towards developed countries as explained by the classic growth model

11 Ravallion's definition was explicitly developed for developing countries (Birdsall, 2010).

(Romer, 2012: p 179). This finding opened a new field in the growth literature based on polarisation analysis (Wang and Tsui, 2000). Esteban and Ray (1994) provided an early fundamental theory on the analysis and measurement of multi-modal polarisation from which some of the indices discussed in this section are based.

The bipolarisation approach estimates the degree of distributional bimodality around the median, giving an intuitive approximation to the size of the middle class and its evolution. Figure II.1 shows two examples of distributions to illustrate this approach. On the left panel, the distribution has two modes around the median, i.e. it is bipolar. By contrast, the distribution on the right panel exhibits only one mode over the median. Considering the middle class as the social group located at the middle of the distribution, a bipolar distribution would suggest a smaller middle class while a unimodal distribution would suggest a larger middle class.

Figure II.1 Examples of bipolar and unimodal distributions



This approach provides a relative measure of the size of the middle class, in which the benchmark of minimum bipolarisation coincides with a situation of perfect equality which, in turn, represents the most “middle-classy” distribution. On the other extreme, maximum *relative* bipolarisation is attained when the bottom half of the population has zero incomes and the top half has an equal strictly positive amount (Yalonetzky, 2016).

Chakravarty (2015: p 38) defines five basic axioms which a rigorous bipolarisation index must fulfil:

- **Decreased spread**: if a progressive transfer between two individuals always on opposite sides of the median renders them closer to each other, then the transfer reduces the bipolarisation level.

- **Clustering-increasing**: if a rank-preserving progressive transfer between two individuals on the same side of the median renders them closer to each other, then the transfer increases the bipolarisation level.
- **Principle of population**: if the same constant value replicates the whole population, then the level of bipolarity will not change.
- **Continuity**: the index is a continuous function of the distribution of the wellbeing domain.
- **Symmetry**: A permutation of incomes among the people in the population should leave the level of bipolarisation unchanged.

Additionally, alternative axioms can detail how the index should react to different possible common transformations of the data. The fulfilment of the scale invariance axiom allows for comparisons between wellbeing domains expressed in different units of income (e.g. pounds, euros, USD).

- **Scale invariance:** if the unit of measurement of a wellbeing domain is multiplied by a constant (e.g. an exchange rate), then the bipolarisation index does not change in value.

Just as in inequality measurement, the imposition of scale invariance means measuring a relative notion of bipolarisation. For the empirical analysis of this chapter, the following bipolarisation indices were selected due to their relevance in the literature: the corrected Foster and Wolfson index and the Wang and Tsui rank-dependent index. Besides, an index based on the area under the Relative Bipolarisation Lorenz curve is introduced as part of the methodological contributions of this work. In the following sections, these three indices are discussed.

II.3.3.1 The corrected Foster-Wolfson index

Foster and Wolfson (1992) proposed the measurement of bipolarisation by partitioning the distribution into two non-overlapping halves through the median and considering the aforementioned axioms. Their Gini-based relative bipolarisation index is:

$$FT = (G_B - G_w)\frac{\mu}{m} \quad \textbf{(II.1)}$$

where G_B represents the between-group Gini index (each half being a group); G_w represents the within-group Gini index; μ is the mean of the whole distribution; and *m* is the respective median.

However, as demonstrated by Rodriguez and Salas (2003), this FT expression could violate the two transfers' axioms of bipolarisation decreasing

spread and cluster increasing even if these transfers do not affect the median. A numerical example can illustrate this effect:

> From an original distribution Y_0=(100, 125, 150, 200, 225, 250), clear-cut mean and media-preserving transfers lead to a new distribution Y_1=(125, 125, 125, 225, 225, 225). These transfers yield a decrease in inequality within polar groups, and an unexpected increase in polarisation according to the original FT index.

Therefore, a modified version of this index without the standardisation by the median was proposed by Rodriguez and Salas (2003) as an adjusted (extended) FT index, which identifies polar groups separated by the mean of the distribution rather than by the median and is defined as follows:

$$FT_{adj} = 2(G_B - G_w) \quad \textbf{(II.2)}$$

This adjusted FT index obtains similar results to the original FT index because both estimate the same G_B when the median and mean converge. In the case of the numerical example mentioned above, the adjusted FT index computes a reduction in bipolarisation rather than an increase.

II.3.3.2 Wang and Tsui

Wang and Tsui (2000) proposed four classes of bipolarisation indices; one of which is a class of rank-dependent indices which could be represented as:

$$WT = \frac{1}{N}\left[\sum_{i=1}^{\frac{N}{2}} a_i y_i^L + \sum_{i=1}^{\frac{N}{2}} b_i y_i^U\right] \quad \textbf{(II.3)}$$

where the individual values of a wellbeing domain, denoted as y, are ordered in the following way: $y_{\frac{N}{2}}^L \leq y_{\frac{N}{2}-1}^L \leq \cdots \leq y_2^L \leq y_1^L \leq m \leq y_1^U \leq y_2^U \leq \cdots \leq y_{\frac{N}{2}-1}^U \leq y_{\frac{N}{2}}^U$.

Wang and Tsui show that equation (II.3) satisfies the two transfer axioms (decreasing spread and clustering increasing) if and only if the weights a_i and b_i fulfil the following inequality: $a_1 < a_2 < \cdots < a_{\frac{N}{2}} < 0 < b_{\frac{N}{2}} < b_{\frac{N}{2}-1} < \cdots < b_2 < b_1$.

Moreover, in order for any index in (II.3) to be equal to a minimum value of zero, only when perfectly equality holds, it must be the case that:

$$\sum_{i=1}^{\frac{N}{2}} a_i + \sum_{i=1}^{\frac{N}{2}} b_i = 0 \quad \textbf{(II.4)}$$

Finally, to have any index in (II.3) to reach a maximum value of 1 only whenever maximum relative bipolarisation holds, then it must be the case that:

$$\sum_{i=1}^{\frac{N}{2}} b_i = \frac{N}{\mu_U} \quad \textbf{(II.5)}$$

and:

$$\sum_{i=1}^{\frac{N}{2}} a_i = -\frac{N}{\mu_U} \quad \forall\ i \in \left[1, \frac{N}{2}\right] \quad \textbf{(II.6)}$$

Even with the above constraints, there are plenty of sensible choices for a_i and b_i. Due to its simplicity for empirical computation (Yalonetzky, 2017), the option chosen for this chapter is the following:

$$b_i = \frac{4\left[\frac{N}{2} - i + 1\right]}{\mu_U\left[\frac{N}{2} + 1\right]} = -a_i \quad \textbf{(II.7)}$$

II.3.3.3 The Relative Bipolarisation Lorenz curve

Yalonetzky (2014) proposed the Relative Bipolarisation Lorenz (RBL) curve. This curve allows for the comparison of different distributions, ranked robustly regarding their relative bipolarisation, using some of the intuition behind Lorenz-type curves and stochastic dominance (Cowell, 2013: p 21). Its mathematical expression is:

$$\psi(p) = \frac{\int_0^p [y_U(q) - y_L(q)]dq}{\mu};\ p \in [0, .5] \quad \textbf{(II.8)}$$

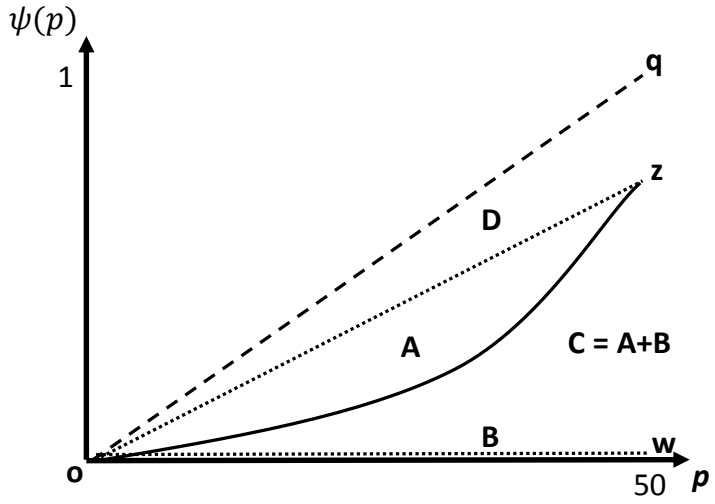
where y_U and y_L represent the income levels of the population above and below the median, respectively, q a determined income percentile, and μ the mean of the distribution. The vector of incomes is ordered so that: $0 \leq y_L(.5) \leq \cdots \leq y_L(0) \leq m \leq y_U(0) \leq \cdots \leq y_U(.5)$, where $y_L(.5)$ and $y_U(.5)$ are the lowest and highest income in the distribution, respectively.

Then, this curve represents the differences between the mean of the p most deprived individuals located above the median and the mean of the p most affluent individuals below the median, standardised by the total mean. Note also that $\psi(.5)$ is equal to G_B multiplied by a constant.

Figure II.2 shows the different possible shapes that the RBL curve could take. The dotted o-w line overlapping with the horizontal axis represents the level of minimum bipolarity or perfect equality; o-q is a dashed line of 45 degrees which shows a situation of maximum bipolarisation. The straight, dotted o-z line represents a situation of perfect bimodality but not the maximum bipolarity benchmark. Finally, any situation between perfect equality and maximum

bipolarisation, except perfect bimodality, is represented by the solid convex curve o-z.

Figure II.2 The Relative Bipolarisation Lorenz (RBL) curve



Source: own representation following Yalonetzky (2014).

A methodological contribution of this chapter on the bipolarisation field is the identification and measurement of four areas observed in the Figure II.2 which can provide valuable information on the dynamics of the middle class. When two distributions' RBL curves do not cross, then a bipolarisation comparison is robust to alternative choices of relative bipolarisation indices within a broad class satisfying the aforementioned axioms (including scale invariance). However, the curve does not directly provide an index to estimate the degree of bipolarisation in a given distribution. In other words, if two RBL curves do not cross, the degree of bipolarisation of one distribution compared to another distribution is unknown. For that purpose, a cardinal index must be chosen (and the result will depend on that choice).

Since potentially several pairs of RBL curves do cross, the bipolarisation ordering of distributions provided by these curves is expected to be only partial. By contrast, any bipolarisation index is capable of ranking all distributions, thereby providing a complete ordering. (By analogy, the Lorenz curve against inequality indices fulfilling the same axioms).

This method has the advantage of analysing bipolarisation under the same conceptual framework as inequality (Rodriguez and Salas, 2003). Then, inspired by the Lorenz curve and the Gini coefficient, and following Euclidean geometry, a definition of four areas below the curves is proposed as follows:

- **Area B.** Consists of the area below the solid convex curve and could be interpreted as the Gini index for bipolarisation[12]. This area can be estimated as:

$$B = \int_0^{.5} \psi(p)dp \quad \textbf{(II.9)}$$

A larger value of B implies an increase in bipolarisation, and therefore a reduction of the middle class; whereas a smaller value implies a reduction in bipolarisation and an increase in the middle class.

- **Area A.** Contained between the dotted line oz and the solid curve oz. This area could be calculated as follows:

$$A = C - B \quad \textbf{(II.10)}$$

This area could be interpreted as the gain or loss in bipolarisation, as the difference between the hypothetical situation of bipolarisation in the absence of any intra-group inequality, and the observed level of bipolarisation.

- **Area C.** Defined as the area below the o-z line, it shows the maximum level of relative bipolarisation given by the upper and lower averages in the absence of intra-group inequality.

The area C could be approximated using Euclidian geometry as the area of a triangle:

$$C = \frac{.5 * z}{2} \quad \textbf{(II.11)}$$

- **Area D.** This area could be estimated as the difference between the area below the straight line o-q and the area C.

$$D = \frac{.5 * q}{2} - C \quad \textbf{(II.12)}$$

However, considering that q is always equal to 1 (because this is the maximum value possible), then the area below o-q is a constant of value .25, rewriting (II.12) as:

$$D = .25 - C \quad \textbf{(II.13)}$$

Area D shows that, for the same total average, the lower half of the distribution has an average greater than zero.

---

12 This interpretation could seem counter-intuitive because, in the original Lorenz curve, oz is a 45-degree line which represents perfect equality and, therefore, A/A+B commonly represents the Gini index. However, in the case of the RBL, ow is the line which represents perfect equality and, therefore, B is the RBL equivalent to the Gini index.

Finally, consider the analysis between different distributions, which could allow for international and inter-temporal comparisons. Let the vertical difference between two independent RBL curves at a percentile i be given by:

$$Diff_i = RBL_{t1i} - RBL_{t2i} \quad \textbf{(II.14)}$$

where RBL_{t1i} and RBL_{t2i} represent values of an RBL curve in period 1 and 2, respectively, at the "*i-th*" percentile. Then Yalonetzky (2014) shows that if and only if $Diff_i \leq 0$ for all i (with at least one strict inequality) then RBL_{t2} will be more bipolarized and have a lower middle class than RBL_{t1}, according to a broad class of relative bipolarisation indices satisfying all the key axioms of symmetry, population principle, scale invariance, decreasing spread, cluster increasing, and normalization. This condition could be defined as stochastic dominance of RBL_{t2} on RBL_{t1}.

If this condition holds with $Diff_i \geq 0$ for all i, then RBL_{t1} will be more bipolarized and has lower middle class compared to RBL_{t2}, fulfilling the condition of stochastic dominance. Otherwise, if two RBL curves cross, then the comparison will not be robust to the choice of alternative indices satisfying the same key axioms (e.g. partial ordering), in a similar way when two traditional Lorenz curves cross (Cowell, 2013: p 35).

II.3.3.4 Concluding remarks on bipolarisation analysis

An alternative to bipolarisation analysis is the multi-modal polarisation (Chakravarty, 2015: p vii). Esteban and Ray (1994) propose variations to their index for distributions with 3, 4 or more mass points (modes). Chakravarty (2015: p 53) shows multipolar indices for discrete and continuous distributions, based on Esteban and Ray (1994). Anderson (2004) suggests undertaking the identification of these modes by cutting the distribution into sub-groups in order to avoid those polarisation trends.

However, Chakravarty affirms that polarisation is observed as "the shrinkage of the middle class" (Chakravarty, 2015: p vii). On the other hand, from a pragmatic perspective, the bipolarisation analysis gives a more precise approximation of the middle class or the absence of it; whereas the inference of the middle class based on multiple modes could be challenging to interpret. Therefore, the analysis of middle class proposed in this work stays on the bipolar side.

II.4 The Latin American context

In this section, a brief description of the Latin American economic context is presented, focusing on the six countries of analysis (Brazil, Chile, Mexico, Paraguay, Peru, and Uruguay). The selection of these sample of countries is based on the data availability required for the estimation of the middle class which will be detailed in the following section.

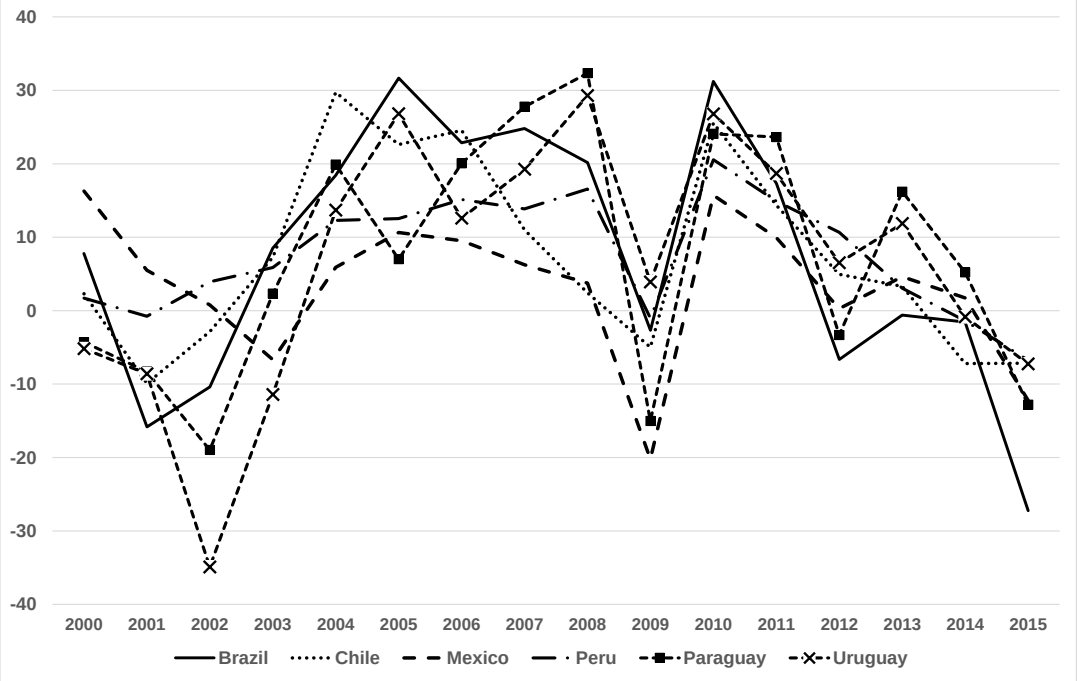
The Latin American region was experiencing a tremendous economic boost during the 2000s until the autumn of 2008 when the effects of the US financial sector collapse spread worldwide. In general terms, the financial crash of 2008 contracted the world economy 6.0% between 2007 and 2009, global unemployment rose, and the world became 10.0% poorer than before the crisis (Temin and Vines, 2013: p 1). The size of the 2008 financial crash, and its consequences, had no comparison in recent history, except for the financial crash of 1929 and its subsequent economic recession (Varoufakis, 2013).

The Latin America economies were affected by this crisis through the progressive economic deceleration of the USA[13]. Thus, the effects of the financial crisis were not felt in the region immediately after the collapse; rather they were delayed. This lagged effects on the regional economy are worth keeping in mind when analysing the regional economic dynamics during this period.

Figure II.3 shows the evolution in the GDP per capita growth rate between 2000 and 2015 for the sample of countries. Trends show three moments of an economic slowdown in this period, with growth on negative figures for some countries: 2002, 2009 and 2015. Whereas in 2009 and 2014 all countries where affected in the same way, during 2002 countries experienced different dynamics and moments of recovery. Despite an immediate recovery from the 2008 shock, all countries experienced a constant and robust slowdown from 2009 onwards, and a deep fall in 2015, suggesting delayed effects of the first 2008 crisis.

13 The later deceleration of China might also have contributed through a fall in export commodity prices.

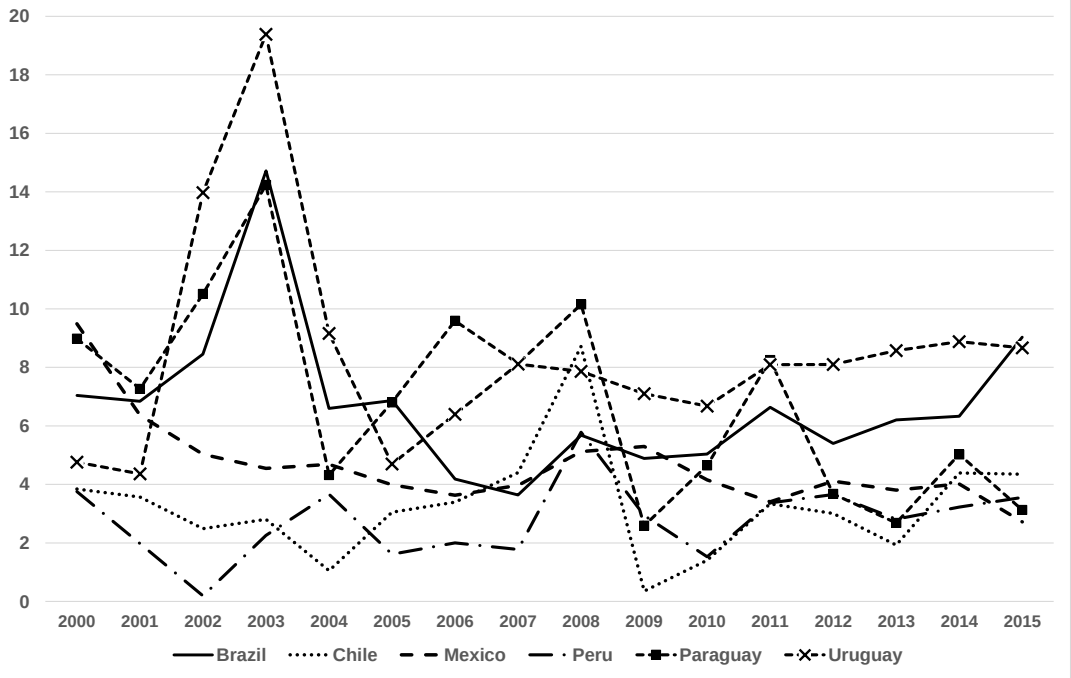
Figure II.3 GDP per capita growth rate, 2000-2015



Source: World Development Indicators, the World Bank, 2015.

Consumption prices exhibit dissimilar evolution across the Latin American countries. However, trends suggest similar patterns among groups of countries. Figure II.4 shows the percent variation on the level of consumer prices between 2000 and 2015. In general, inflation presents a variation below 10% in the region. Uruguay, Paraguay and Brazil had the highest variations in inflation during this period, particularly in 2003, as well as a path of close price relationship among them. Mexico, on the other hand, has a clear trend of inflation reduction; meanwhile Chile and Peru show a similar stable inflation path with a shock in 2008, possibly due to the variation in commodities prices product of the financial crisis.

Figure II.4 Consumer price index, percent variation, 2000-2015

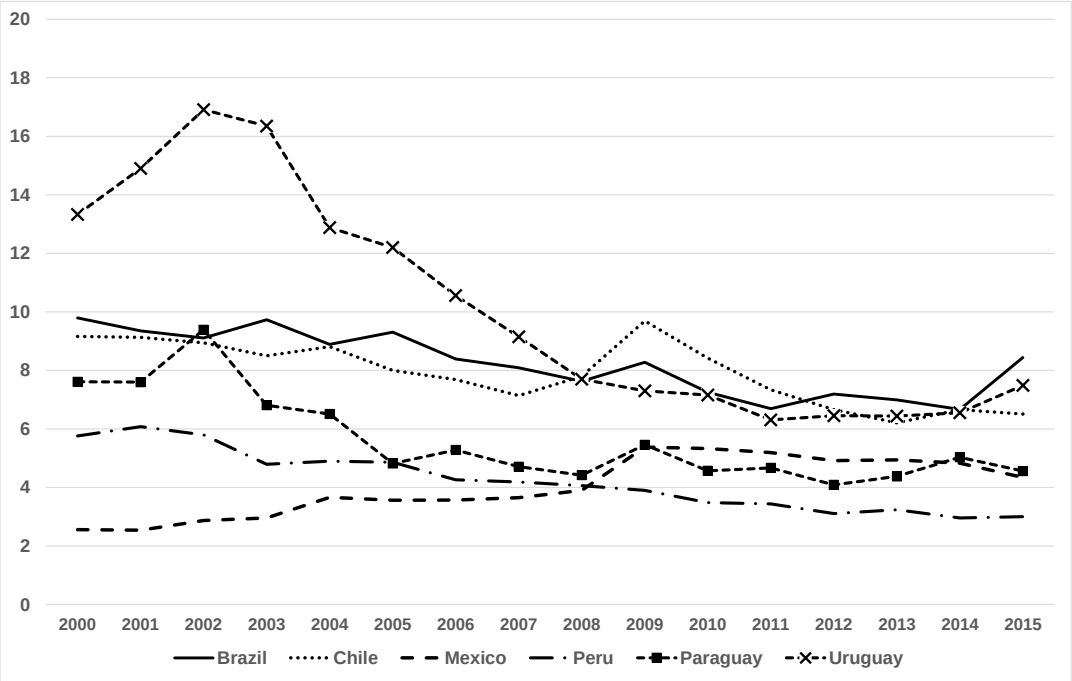


Source: World Development Indicators, the World Bank, 2015.

Labour is the most important source of income in Latin American households, representing up to 80% of their total incomes (Abeles et al., 2014). Considering the lack of strong welfare states in the region, employment status becomes one of the most critical survival conditions for the population. Further analysis on employment and middle class, particularly for women in Mexico, will be conducted in chapter IV.

Figure II.5 shows the evolution of the unemployment rate for the six Latin American countries selected to analysis in the period 2000 to 2015. On average, all countries are experiencing unemployment reductions, except for Mexico where unemployment increased after 2008 and maintains higher levels than before 2008, possibly explained by the high economic integration with the USA. The 2008 financial crisis affected employment negatively in all these countries except for Uruguay and Peru. Uruguay had a major shock on unemployment in 2006, reaching levels higher than 10%. In general, unemployment rates are below 10% in these countries, and a few permanent effects from the 2008 financial crisis could be seen.

Figure II.5 Unemployment rate, 2000-2015



Source: World Development Indicators, the World Bank, 2015.

The financial crisis of 2008 had a profound negative influence on the world's economy, generating economic recession, uncertainty and unemployment. Latin America has not been the exception. However, due to the characteristics of this crisis, its impact has not been felt immediately in this region but delayed, following the growth path of leading economies and partners like the US and China. In general terms, the region has experienced a slowdown of economic activity, without recovering its strength to pre-crisis levels. Eight years after the financial crisis, the economy in the Latin American region is not recovering; on the contrary, it seems that the region has entered into a period of economic recession.

Moreover, not all the countries were affected in the same way. Countries more integrated to the US economy suffered the worst falls in productivity and rise of unemployment immediately after the economic shock occurred but recovering relatively faster than other countries. On the other hand, economies less integrated to the US but more integrated among them at regional level, or more dependent on China, present a delay in the negative impact of the crisis; e.g. Brazil, Chile and Uruguay. In conclusion, a variegated evolution of the middle class should be expected across countries within the period of analysis.

II.5 Empirical analysis

This section presents the empirical implementation of the middle-class definitions, described in section II.3. This statistical exercise depends on the availability of household or individual micro-data[14]. Information aggregated at national level like balance sheets or national accounts is not useful for this case because it hides the real distribution of the wellbeing domain among individuals.

Thus, only six countries in the region offered complete and publicly available micro-datasets. These countries are Brazil, Chile, Mexico, Paraguay, Peru, and Uruguay. Not all these surveys contain information on household expenditure; besides, data collection methods for expenditure vary among countries. Therefore, this analysis focuses only on the wellbeing domain of income[15].

The reason for choosing countries in Latin American instead of other regions is to provide a general framework to contextualise and embed further analyses on the Mexican middle class[16]. Latin American countries share cultural backgrounds and historical roots, reinforced by a common language (or linguistic proximity in the case of Brazil), similar levels of economic development, and regional economic integration[17].

II.5.1 Data

To guarantee consistency in the estimations of the middle class, homogeneous series of the wellbeing domain are required at household level for comparability, i.e. without any significant methodological changes on data collection over time. In Latin American, this type of statistical information is not always available for all countries, limiting the outreach of this analysis.

As mentioned before, six Latin American countries were selected for this study: Brazil, Chile, Mexico, Paraguay, Peru and Uruguay. Their household

---

14 According to the wellbeing domain of interest. Income could be reported at individual or household level, whereas expenditure is exclusively reported at household level.

15 Details on data availability will be discussed in subsection II.5.1.

16 These analyses correspond to chapters III, IV and V of this thesis.

17 Developed countries were not considered either in this analysis in order to guarantee homogeneity among the sample.

surveys offer detailed and consistent information on income, statistical representativeness of urban and rural areas, and enough sample size over time. Moreover, the micro-datasets are free to download. Table II.3 shows a general description of these surveys by country.

Table II.3 Summary of household surveys

Country	Survey	Organization	Period available[1]	Frequency	Income	Expenditure
Brazil	National Household Sample Survey (PNAD)	Brazilian Institute for Geography and Statistics (IBGE)	2004-2015	Annual	Yes	Yes
Chile	National Socio-economic Characterisation Survey (CASEN)	Ministry of Social Development (MSD)	2000-2011	Triennial	Yes	No
Mexico	National Incomes and Expenditures Households Survey (ENIGH)	National Institute for Geography and Statistics (INEGI)	2000-2014	Biannual	Yes	Yes
Paraguay	Permanent Households Survey (EPH)	General Directorate of Statistics, Surveys and Censuses (DGEEC)	2002-2016	Annual	Yes	No
Peru	National Households Survey (ENAHO)	National Institute for Statistics and Informatics (INEI)	2004-2015	Annual	Yes	Yes
Uruguay	Households Continue Survey (ECH)	National Institute for Statistics (INE)	2000-2017	Annual	Yes	No

Source: own construction based on information from each statistical office.
[1] Refers to the period where income information is comparable between all years.

The Brazilian IBGE publishes the PNAD survey yearly since 2004, being 2015 the latest available, except for 2010 due to overlapping with the Census. This survey has an average sample size of 113,203 households each year. For Chile, the MSD publishes the CASEN survey every three years since 2000, being 2011 the latest comparable sample and an average sample size of 72,967 households per year.

The Mexican INEGI publishes the ENIGH survey every two years, except 2005, with comparable information from 2000 to 2014. The sample size grew from over 10,000 households in 2000 to more than 27,000 in 2010. However, the sample decreased to 9,000 in 2012 due to adjustments in INEGI's surveys, without affecting comparability with previous years. Paraguay's DGEEC publishes the EPH survey, which has yearly comparable information from 2002 to 2016, and an average sample of 5,630 households per year.

The Peruvian ENAHO, published by the INEI, provides comparable information on households' characteristics every year from 2004 to 2015, giving an average sample size of 23,466 households each year. Finally, the Uruguayan INE publishes the ECH survey yearly, with comparable information from 2005 to 2017, and average sample size is of 50,944 households per year. Appendix A.1 shows detailed information on sample sizes and descriptive statistics for income for each country and year.

As shown in table II.3, not all surveys provide information on household expenditure. Moreover, the methods to collect household expenditure vary among them. In the case of the Mexican ENIGH, and to some extent in the Peruvian ENAHO, a vast set of questions collect detailed information on

expenditure and frequency of items purchase, whereas in the Brazilian PNAD a single question captures an estimated household monthly expenditure.

The collection methods for income information are similar among surveys. All of them identify primary income sources like labour income, rents and profits, transfers from social development and government programs, scholarships, remittances from abroad, and transfers between households. Some surveys collect information in a more exhaustive way than others, but in general, all these household income sources are covered. Therefore, income is selected for analysis rather than expenditure.

The household income information is comparable and consistent within the countries in the period of analysis, but it is not comparable between countries because some surveys report total income whereas others report current income[18]. This is the reason why this study follows the evolution of the middle class within countries and do not compare the evolution between countries.

All these surveys provide, besides monthly household income, the total number of household members. Thus, the total monthly per capita income is computed as follows:

$$Total\ Monthly\ per\ Capita\ Income_i = \frac{Total\ Monthly\ Income_i}{Household\ Size_i} \quad \textbf{(II.15)}$$

where the total monthly income for household *i* is divided by the total number of members in household *i*. Even though data is adjusted for inflation in several surveys, nominal income could be used directly if necessary, instead of deflated values, thanks to the scale invariance property satisfied by all measures of bipolarisation used in this chapter.

A per capita income based on the absolute value of the household size does not take into consideration differences in needs between household members (Atkinson et al., 1994), or economies of scale (Forster, 1994). An alternative is to characterise household members according to a particular value equivalent to their needs.

In this way, the OECD gives a value of 1 to the household head, 0.5 to any extra adult in the household, and 0.3 to each child (OECD, 2019). In Mexico, the estimation of poverty gives a value of 1 to the household head, 0.99 to any extra adult in the household, 0.71 to each teenager, 0.74 to each child

---

18 Total income represents the sum of all types of income a household receives in a certain period. Current income represents the sum of only recurrent-types of income, not considering extraordinary, once in a life, incomes.

between 6 to 12 years old, and 0.70 to each child between 0 to 5 years old (CONEVAL, 2018). Thus, the equivalent per capita income would be larger than an absolute per capita income.

However, an equivalent per capita income is not possible to calculate in this analysis due to limitations to information. The number of household members is available in all surveys, but not always their ages. The implications of not using an equivalent per capita income are possible biases on the middle class estimations towards the concentration of populous households (on the deprived or on the affluent side).

Considering that most countries have comparable data series since early 2000, the analysis timeframe ranges from 2000 to 2015, depending on the country data availability and periodicity. In the following subsection, estimation results for all the different middle-class definitions are presented.

II.5.2 Results

In this subsection, results on the different middle-class definitions, discussed in section II.3, are presented. First, this section shows estimations for relative definitions, followed by absolute definitions, and finally for the analysis based on relative bipolarisation. Estimations were calculated for each country mentioned before, during the period where comparable data is available. General information on the relative and absolute thresholds is shown in tables II.1 and II.2 respectively[19].

II.5.2.1 Results based on relative definitions

As mentioned in section II.3.1, relative definitions of the middle class could be divided into two categories: estimations based on percentages of the median, and estimations based on population percentiles as defined by Levy (1987). For the former category, table I.4 shows estimation for all countries considering thresholds at 50% and 150% of the median (MC-1), at 75% and 125% (MC-2), and 60% and 225% of the median (MC-3).

---

19 The particular threshold values for each definition, country and year is shown in Appendix A.2.

Table II.4 Percentage of the population identified as middle class based on relative definitions, for all countries and years available

Year	Brazil			Chile			Mexico			Paraguay			Peru			Uruguay		
	MC-1	MC-2	MC-3	MC-1	MC-2	MC-3	MC-1	MC-2	MC-3	MC-1	MC-2	MC-3	MC-1	MC-2	MC-3	MC-1	MC-2	MC-3
2000				40.8%	20.3%	47.1%	46.2%	22.5%	53.2%									
2001																		
2002							47.4%	24.0%	55.0%	39.7%	18.9%	47.7%						
2003				50.1%	24.7%	56.3%				44.0%	21.7%	51.0%						
2004	43.4%	22.3%	47.9%				46.5%	23.1%	53.2%	46.1%	23.0%	53.0%	34.4%	17.0%	43.3%			
2005	43.9%	22.1%	48.4%				44.5%	22.1%	51.6%	43.9%	21.8%	51.7%	33.0%	15.8%	41.7%			
2006	44.3%	23.4%	49.9%	52.3%	26.2%	58.2%	46.2%	22.9%	53.4%	44.2%	23.2%	52.4%	33.4%	16.3%	42.2%	52.4%	26.9%	60.6%
2007	45.0%	23.1%	50.1%							45.7%	22.9%	52.2%	33.4%	16.6%	42.4%	51.1%	25.7%	58.6%
2008	46.1%	23.7%	51.5%				47.1%	23.4%	54.1%	45.5%	22.6%	52.4%	34.3%	17.0%	43.7%	53.4%	27.6%	61.0%
2009	46.3%	23.6%	51.6%	54.6%	28.8%	61.9%				43.3%	21.1%	51.4%	35.4%	17.7%	44.5%	53.1%	26.6%	60.3%
2010							48.0%	23.8%	55.2%	45.2%	22.9%	52.8%	36.7%	18.5%	46.0%	54.3%	27.6%	61.9%
2011	48.0%	24.7%	53.1%	51.5%	26.1%	57.8%				45.9%	23.5%	53.7%	37.4%	18.9%	47.3%	56.3%	29.2%	64.1%
2012	48.0%	25.8%	53.8%				48.7%	23.8%	55.1%	46.9%	24.0%	55.1%	38.1%	19.4%	48.2%	56.7%	29.3%	64.8%
2013	48.7%	25.1%	53.9%							47.0%	23.3%	55.6%	38.0%	19.0%	47.1%	58.0%	30.6%	66.4%
2014	49.1%	26.1%	55.5%				51.3%	25.9%	58.0%	48.7%	24.0%	56.1%	38.5%	19.3%	47.2%	58.5%	30.8%	67.2%
2015	49.9%	26.3%	55.2%							46.5%	23.6%	55.1%	39.6%	19.4%	47.9%	57.9%	29.9%	66.2%

Source: own estimations using microdata from: Brazil, PNAD various years; Chile, PNAD various years; Mexico, ENIGH various years; Paraguay, EPH various years; Peru, ENAHO various years; Uruguay, ECH various years.

These results show a general and consistent increase in the size of the middle class, as a percentage of the total population, between all relative definitions based on percentages of the median in the period of analysis. Considering that the medians differ across countries, Chile has a more considerable increase in the middle class in all the indices, whereas Peru has the smallest changes. Even a detailed look at the variations between years suggests consistency on marginal increases or decreases. Definition MC-3 has the wider thresholds and, therefore, the higher incidences on the middle class, whereas MC-2 has lower thresholds and lower incidences.

Table II.5 shows estimations regarding relative definitions based on percentiles of per capita income. These middle-class definitions follow the discussion in subsection II.3.1, considering thresholds as population belonging to a certain income percentile. These definitions are: percentiles 20 to 80 (R-1), percentiles 40 to 60 (R-2), and percentiles 20 to 90 (R-3).

Table II.5 Middle-class income ratios based on relative definitions, percentiles, for all countries and years available

Year	Brazil			Chile			Mexico			Paraguay			Peru			Uruguay		
	R-1	R-2	R-3	R-1	R-2	R-3	R-1	R-2	R-3	R-1	R-2	R-3	R-1	R-2	R-3	R-1	R-2	R-3
2000				33.1%	9.9%	48.5%	39.5%	12.2%	55.3%									
2001																		
2002							40.4%	12.5%	56.3%	35.9%	11.0%	51.3%						
2003				36.3%	11.2%	51.0%				34.6%	10.5%	49.1%						
2004	34.5%	10.5%	51.1%				37.6%	11.5%	53.1%	36.5%	11.2%	51.5%	37.5%	11.4%	54.5%			
2005	34.8%	10.5%	50.5%				38.7%	11.8%	54.8%	38.9%	11.8%	55.1%	36.4%	10.9%	53.4%			
2006	35.5%	10.7%	51.1%	38.1%	11.9%	53.0%	39.4%	12.1%	55.6%	36.9%	11.3%	52.2%	36.9%	11.2%	53.9%	45.2%	14.3%	61.2%
2007	36.2%	11.1%	52.0%							35.0%	10.7%	49.5%	36.6%	11.1%	53.6%	44.4%	14.0%	60.7%
2008	36.4%	11.3%	52.0%				38.6%	11.9%	54.2%	37.7%	11.5%	53.2%	37.5%	11.4%	54.5%	45.3%	14.4%	61.3%
2009	36.9%	11.5%	52.4%	41.0%	12.8%	55.9%				39.6%	12.2%	55.8%	37.5%	11.4%	54.0%	45.0%	14.1%	61.1%
2010							40.6%	12.5%	56.8%	37.0%	11.5%	51.9%	39.2%	12.0%	56.0%	45.7%	14.3%	61.8%
2011	38.4%	11.8%	53.6%	38.6%	11.9%	54.0%				39.2%	12.2%	54.5%	39.6%	12.3%	56.0%	46.8%	14.8%	62.8%
2012	39.1%	11.8%	54.4%				40.1%	12.3%	56.1%	41.3%	13.0%	57.0%	40.9%	12.8%	57.6%	47.8%	15.2%	63.8%
2013	38.6%	12.5%	54.0%							41.0%	12.7%	56.6%	40.5%	12.6%	57.6%	48.4%	15.4%	64.3%
2014	39.4%	12.2%	53.9%				40.9%	12.7%	56.7%	40.0%	12.5%	55.2%	40.0%	12.3%	56.9%	48.1%	15.3%	63.8%
2015	39.2%	12.6%	55.4%							42.2%	13.1%	58.3%	40.1%	12.4%	57.0%	47.9%	15.2%	63.7%

Source: own estimations using microdata from: Brazil, PNAD various years; Chile, PNAD various years; Mexico, ENIGH various years; Paraguay, EPH various years; Peru, ENAHO various years; Uruguay, ECH various years.

Because all these definitions are based on a population domain, the percentage of the population will remain constant for all years (the proportion of the population will always be within the same percentile). Therefore, this table reports the income shares of the groups identified as middle class as percentages of total national income.

Results show a consistent increase in the income share of the middle class across all countries and years. Marginal changes, however, are not necessarily consistent, like in the case of Brazil, where R-1 and R-3 decreased between 2012 and 2013, but R-2 increased. As expected from results of relative definitions, wider thresholds lead to larger proportions of middle-class income.

II.5.2.2 Results based on absolute definitions

Following subsection II.3.2, four definitions are reported with thresholds as follow: $2 to $13 USD PPP per day (MC-4); $10 to $50 USD PPP per day (MC-5); $2 to $10 USD PPP per day (MC-6); and $12 to $50 USD PPP per day (MC-7). PPP values are estimated based on World Bank estimations (World Bank, 2018). To get per day income values, the monthly per capita income from equation (II.15) was divided by 30.4, the average number of days per month. Table II.6 shows the estimated proportion of population identified as middle class for these absolute definitions.

In this case, an important discrepancy on the trends between middle-class definitions is observed, opposite to the general consistency between trends of relative definitions. Definitions MC-4 and MC-6 show a decrease in the middle class in estimations for all countries. Definitions MC-5 and MC-7, on the other hand, show an increase in the middle-class incidence for all countries. These results highlight the relevance of the threshold definitions and build on the controversy of the arbitrary threshold selection.

Table II.6 Percentage of the population identified as middle class based on absolute definitions, for all countries and years available

Year	Brazil				Chile				Mexico			
	MC-4	MC-5	MC-6	MC-7	MC-4	MC-5	MC-6	MC-7	MC-4	MC-5	MC-6	MC-7
2000					61.4%	18.5%	54.9%	13.5%	68.6%	29.5%	59.5%	22.8%
2001												
2002									67.9%	32.1%	57.8%	24.9%
2003					67.5%	34.8%	56.6%	26.8%				
2004	61.2%	34.4%	51.8%	28.3%					62.3%	39.1%	51.0%	30.9%
2005	59.6%	36.5%	50.2%	30.1%					61.5%	38.5%	51.1%	31.3%
2006	58.2%	42.9%	44.5%	31.6%	62.4%	43.1%	49.4%	33.9%	59.8%	42.6%	48.5%	34.6%
2007	56.4%	45.1%	42.0%	33.9%								
2008	54.5%	47.5%	40.2%	36.4%					58.1%	45.6%	46.3%	37.1%
2009	54.2%	48.0%	39.3%	37.0%	56.8%	50.5%	42.1%	40.0%				
2010									60.0%	43.8%	48.1%	35.3%
2011	51.1%	51.7%	36.1%	44.7%	43.4%	59.7%	29.5%	50.1%				
2012	43.4%	54.1%	33.5%	47.2%					60.0%	44.4%	48.3%	36.2%
2013	41.2%	54.6%	32.4%	48.1%								
2014	40.3%	57.0%	31.0%	50.5%					58.3%	48.2%	45.4%	38.9%
2015	41.6%	56.2%	32.0%	49.2%								
Year	**Paraguay**				**Peru**				**Uruguay**			
	MC-4	MC-5	MC-6	MC-7	MC-4	MC-5	MC-6	MC-7	MC-4	MC-5	MC-6	MC-7
2000												
2001												
2002	62.0%	30.9%	52.9%	24.5%								
2003	65.8%	29.2%	57.0%	23.0%								
2004	67.9%	29.7%	58.9%	23.4%	36.1%	50.2%	28.2%	44.9%				
2005	62.7%	35.4%	53.4%	28.9%	37.1%	48.3%	29.6%	43.2%				
2006	64.5%	34.6%	53.9%	27.3%	34.3%	49.9%	26.9%	44.8%	52.5%	56.4%	38.4%	46.5%
2007	61.9%	38.2%	51.5%	30.5%	30.6%	49.7%	24.2%	45.5%	42.0%	62.5%	29.6%	53.9%
2008	62.4%	39.9%	51.2%	31.8%	27.6%	50.3%	21.5%	46.2%	35.2%	67.9%	22.9%	59.6%
2009	59.5%	40.7%	48.9%	33.8%	26.3%	51.0%	20.1%	46.8%	31.4%	68.9%	19.7%	61.1%
2010	56.7%	45.7%	44.7%	37.3%	24.7%	51.6%	19.0%	47.5%	27.1%	71.7%	16.2%	64.6%
2011	53.7%	49.2%	41.2%	40.5%	23.5%	51.7%	18.1%	48.2%	23.1%	74.5%	12.9%	67.8%
2012	51.4%	52.7%	38.6%	43.5%	21.5%	51.0%	15.9%	47.3%	19.7%	75.3%	10.8%	69.4%
2013	46.2%	56.3%	34.5%	48.0%	20.6%	49.6%	15.1%	45.8%	17.1%	75.8%	9.3%	70.7%
2014	46.2%	57.6%	33.8%	48.7%	19.6%	50.3%	13.9%	46.5%	14.4%	75.7%	7.8%	71.5%
2015	45.9%	57.4%	34.1%	49.4%	19.0%	51.4%	13.4%	47.5%	14.5%	74.5%	8.0%	70.3%

Source: own estimations using microdata from: Brazil, PNAD various years; Chile, PNAD various years; Mexico, ENIGH various years; Paraguay, EPH various years; Peru, ENAHO various years; Uruguay, ECH various years.

II.5.2.3 Results based on bipolarisation analysis

Following subsection II.3.3, the bipolarisation indices estimated were the Wang-Tsui (WT), the corrected Foster-Wolfson (FW) from equation (II.2), and the four areas of bipolarisation represented in the figure (II.2), A, B, C and D, following equations (II.9), (II.10), (II.11), and (II.13) respectively. The point z, i.e. a multiple of the between-group Gini coefficient G_B, is reported for reference. Table II.7 shows the estimations for all bipolarisation approaches.

Table II.7 Evolution of bipolarisation by different definitions

Year	Brazil							Chile						
	WT	FW	Z	A	B	C	D	WT	FW	Z	A	B	C	D
2000								0.327	0.143	0.957	0.872	0.086	0.957	0.043
2001														
2002														
2003								0.308	0.128	0.917	0.818	0.100	0.917	0.083
2004	0.334	0.141	0.873	0.710	0.164	0.873	0.127							
2005	0.333	0.141	0.867	0.701	0.166	0.867	0.133							
2006	0.333	0.140	0.871	0.713	0.158	0.871	0.129	0.310	0.127	0.878	0.758	0.120	0.878	0.122
2007	0.333	0.138	0.841	0.658	0.183	0.841	0.159							
2008	0.327	0.136	0.854	0.692	0.163	0.854	0.146							
2009	0.330	0.137	0.856	0.695	0.161	0.856	0.144	0.311	0.124	0.771	0.599	0.172	0.771	0.229
2010														
2011	0.325	0.133	0.807	0.618	0.189	0.807	0.193	0.320	0.131	0.867	0.736	0.131	0.867	0.133
2012	0.324	0.132	0.807	0.622	0.185	0.807	0.193							
2013	0.324	0.131	0.798	0.607	0.191	0.798	0.202							
2014	0.321	0.132	0.792	0.605	0.187	0.792	0.208							
2015	0.320	0.131	0.766	0.562	0.204	0.766	0.234							

Year	Mexico							Paraguay						
	WT	FW	Z	A	B	C	D	WT	FW	Z	A	B	C	D
2000	0.346	0.144	0.755	0.523	0.232	0.755	0.245							
2001														
2002	0.346	0.143	0.915	0.812	0.103	0.915	0.085	0.343	0.149	0.878	0.715	0.163	0.878	0.122
2003								0.315	0.135	0.912	0.794	0.118	0.912	0.088
2004	0.331	0.139	0.928	0.830	0.098	0.928	0.072	0.320	0.135	0.851	0.692	0.159	0.851	0.149
2005	0.350	0.147	0.836	0.657	0.179	0.836	0.164	0.356	0.150	0.748	0.490	0.259	0.748	0.252
2006	0.349	0.145	0.762	0.533	0.229	0.762	0.238	0.328	0.139	0.920	0.814	0.106	0.920	0.080
2007								0.306	0.131	0.919	0.812	0.107	0.919	0.081
2008	0.335	0.140	0.927	0.830	0.096	0.927	0.073	0.338	0.142	0.803	0.601	0.203	0.803	0.197
2009								0.361	0.152	0.730	0.456	0.274	0.730	0.270
2010	0.349	0.143	0.767	0.553	0.213	0.767	0.233	0.321	0.136	0.872	0.729	0.143	0.872	0.128
2011								0.334	0.139	0.809	0.622	0.188	0.809	0.191
2012	0.346	0.143	0.776	0.571	0.205	0.776	0.224	0.346	0.143	0.743	0.515	0.228	0.743	0.257
2013								0.346	0.142	0.749	0.523	0.225	0.749	0.251
2014	0.335	0.136	0.783	0.600	0.182	0.783	0.217	0.331	0.137	0.822	0.655	0.167	0.822	0.178
2015								0.357	0.146	0.773	0.566	0.207	0.773	0.227

Year	Peru							Uruguay						
	WT	FW	Z	A	B	C	D	WT	FW	Z	A	B	C	D
2000														
2001														
2002														
2003														
2004	0.391	0.170	0.799	0.587	0.212	0.799	0.201							
2005	0.391	0.171	0.821	0.620	0.201	0.821	0.179							
2006	0.393	0.172	0.765	0.510	0.255	0.765	0.235	0.349	0.137	0.700	0.485	0.215	0.700	0.300
2007	0.389	0.170	0.797	0.565	0.232	0.797	0.203	0.354	0.141	0.720	0.506	0.215	0.720	0.280
2008	0.390	0.170	0.785	0.551	0.234	0.785	0.215	0.346	0.136	0.704	0.494	0.210	0.704	0.296
2009	0.382	0.166	0.757	0.508	0.249	0.757	0.243	0.350	0.138	0.709	0.496	0.213	0.709	0.291
2010	0.387	0.166	0.744	0.501	0.243	0.744	0.256	0.348	0.136	0.689	0.475	0.214	0.689	0.311
2011	0.382	0.163	0.744	0.507	0.237	0.744	0.256	0.343	0.132	0.640	0.415	0.224	0.640	0.360
2012	0.389	0.165	0.770	0.554	0.216	0.770	0.230	0.346	0.132	0.633	0.409	0.224	0.633	0.367
2013	0.393	0.167	0.729	0.478	0.251	0.729	0.271	0.341	0.129	0.576	0.335	0.241	0.576	0.424
2014	0.386	0.164	0.810	0.625	0.185	0.810	0.190	0.335	0.127	0.607	0.384	0.223	0.607	0.393
2015	0.383	0.163	0.714	0.459	0.254	0.714	0.286	0.340	0.129	0.602	0.370	0.232	0.602	0.398

Source: own estimations using microdata from: Brazil, PNAD various years; Chile, PNAD various years; Mexico, ENIGH various years; Paraguay, EPH various years; Peru, ENAHO various years; Uruguay, ECH various years.

The WT and FW indices show a consistent reduction of bipolarisation, hence an increase in the middle class, on the overall period for Brazil, Chile, Mexico, Peru and Uruguay. However, in the case of Paraguay, a discrepancy arises, were WT increase, but FW decreases, leading to an inconclusive change in the middle class.

For the purpose of inferring overall changes in the middle class, area B is the most helpful tool: areas A, C and D measure the closeness to hypothetical conditions of general perfect bimodality or maximum relative bipolarisation,

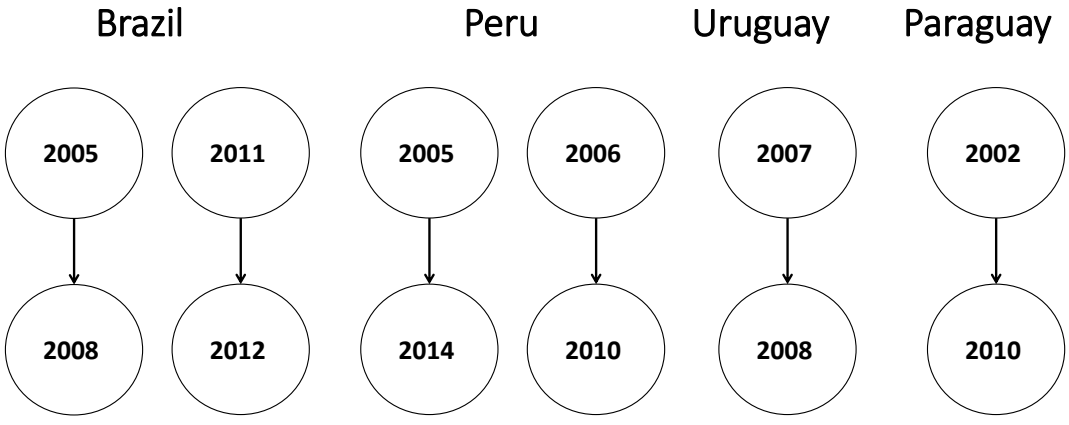
whereas area B shows the distance to a condition of minimum bipolarisation (where every observation is the mode). A smaller B shows less distance to minimum bipolarity and a higher middle class. Appendix A.3 presents all the single RBL curves estimated by country.

As noted in section II.3.3.3, in order to observe changes in the middle class between two periods (periods A and B) using RBL curves, the distribution of one curve has to be larger (dominate) the other curve without intersections. This is a condition of stochastic dominance of order 1. In this example, if A dominates B (A>B), then a higher level of bipolarisation is observed in period A compared to period B, suggesting an expansion of the middle class from period A to B.

An empirical limitation of this approach is that the differences between the RBLs are so small that a graphical visualisation of the curves is impractical. Accordingly, an analysis of stochastic dominance was conducted between all RBL curves computed for each country following equation (II.14), and Hasse diagrams support the graphical analysis.

Hasse diagrams show the relationship of stochastic dominance between distributions as an arrow pointing from the dominant period to the dominated, and could be interpreted as the year with a lesser middle class (more bipolar) pointing to the year with a greater middle class (less bipolar). Figure II.6 shows these Hasse diagrams.

Figure II.6 Hasse diagrams for Brazil, Peru, Uruguay and Paraguay, 2000-2015



Source: own estimations using microdata from: Brazil, PNAD various years; Paraguay, EPH various years; Peru, ENAHO various years; Uruguay, ECH various years.

These diagrams suggest a consistent reduction in bipolarisation for these particular countries and years, meaning increases in their middle class respectively. Only six country-year comparisons, out of hundreds of possibilities, show robust declines in relative bipolarisation (neither of which

includes Chile or Mexico). The absence of any robust decreases in the middle class (as measured by the relative bipolarisation concept) is also noteworthy.

Based on these diagrams, the comparison of the B areas is only valid when the condition of complete ordering is fulfilled (no crossings in RBLs). In the case of Brazil, the areas to compare are B-2005 to B-2008 and B-2011 to B-2012. A reduction on the B area is observed in both periods (0.166-0.163 and 0.189-0.185 respectively), suggesting a reduction in bipolarisation, and an increase in the middle class. These results are also consistent with FW and WT.

In conclusion, even though bipolarisation analysis dispenses with the need to choose arbitrary middle-class thresholds, the empirical analysis shows no consistency among a selection of feasible bipolarisation measurement approaches. These empirical findings raise a warning, and their potential implications should be further explored and debated in further studies.

I.6 Conclusion

This chapter analyses the definition of the middle class in Latin America, following the main motivation of investigating the effects of, and differences between, its measurement approaches (traditional relative and absolute definitions, as well as bipolarisation analysis), from a unidimensional perspective.

Following the literature review of this chapter, the relevance of the middle class for the development of the Latin American region becomes clear. However, the lack of consensus on its definition, and the arbitrariness of the thresholds selection could lead to inconclusive results.

An empirical analysis of the middle-class evolution was conducted, using the per capita household income as wellbeing domain, during the period 2000-2015, for six countries: Brazil, Chile, Mexico, Paraguay, Peru, and Uruguay. This sample of countries was selected due to data availability (income information at household level is required to compute all the middle-class definitions).

These results are not all conclusive about the expansion or shrinkage of the middle class due to estimation discrepancies, which was an expected outcome considering the arbitrariness nature of the middle-class threshold definition as discussed before. Estimations for relative definitions show an

expansion of the middle class for all threshold definitions and all countries. Moreover, income ratios show an increase in the income share of the middle class in the total income for all population percentiles and countries.

The bipolarisation analysis shows a decrease in bipolarity during this period in all the countries in the sample, inferring an expansion in the middle class. However, a discrepancy between FW and WT bipolarisation indices is found in the case of Paraguay, where it is not possible to conclude whether income bipolarisation has increased or reduced.

Estimations based on absolute definitions exhibit discrepancies. Definitions for developing countries (with lower-middle-class thresholds, between $2-$13 USD PPP), show a decrease in the middle class for all countries. On the other hand, definitions for developed countries (with upper-middle-class thresholds, between $10-$50 USD PPP) suggest an expansion in the middle class.

The RBL curve analysis presents a few dominance relationships in all countries, except for Chile and Mexico where no relationships are found. These relationships, however, cover most of the period of analysis, suggesting mid-term changes in bipolarisation (except for Uruguay where the relation covers only one year). In all cases, the dominance relationships suggest a decrease in bipolarisation over time, inferring an increase in the middle class. These results are supported by the analysis on B areas, which shows a decrease in bipolarisation, and therefore an expansion of the middle class.

In conclusion, all these results could be interpreted as an expansion of the middle class in this period, except for absolute middle-class definitions based on low thresholds ($2 to $13 and $2 to $10 USD PPP per day), which show a contraction of this group. A possible explanation to these results would be a translation shift of the income distribution of all these countries toward higher incomes during the whole period, suggesting that the middle class in all countries is better off.

Therefore, based on these results, middle-class definitions for the Latin America region should consider the dispersion of the income distribution. These findings could contribute to overcome the concerns expressed by Daza and Cortes (2013) about the relevance of a middle-class definition for the Latin American context.

Regarding the methodological contributions of this chapter, the advantages of the analysis of RBL areas is the absence of controversial threshold definitions. However, one of its main drawbacks is the absence of direct identification of the individuals who belong to the middle class: this approach provides

information to build an inference on the evolution of the middle class but does not allow to identify particular individuals.

Another current limitation of the bipolarisation analysis is the absence of an analytical method to test the statistical differences between distributions. Future work could develop an analytical test for differences based on current Lorenz curves test for generalised stochastic dominance like Arora and Jain (2006) or Dardanoni and Forciana (1999).

Future research on this topic is related to the interpretation and analysis of the other RBL curve areas calculated in the empirical analysis (A, C, D), as well as the development of a statistical test of differences which could confirm whether the changes in bipolarisation, from the area perspective, are statically significant or not.

Chapter III
The middle class in Mexico

III.1 Introduction

This chapter analyses the middle class as a multidimensional construct, and proposes two methodological approaches for the identification and measurement of the middle class in Mexico, based on counting the household characteristics which define the middle, the deprived and the affluent classes (counting approach). The empirical analysis uses the Mexican survey MCS-ENIGH for 2012.

The first methodological approach consists of selecting a specific set of household characteristics which define each of the social groups, and counts these characteristics in order to identify to which social class the household belongs. Thus, two ways of middle-class identification are developed: direct identification; or indirect identification whereby the deprived and the affluent are identified first, and then those outside these two groups are deemed to belong to the middle class.

The second methodological approach is inspired by Castellani et al. (2015), and identifies the middle class as the population above the multidimensional poverty condition, as defined by the Mexican government, and below the top threshold of a unidimensional income middle-class definition. In this way, the combination of indirect identification and unidimensional middle-class definition makes this method a special case of the counting approach.

The main research question of this chapter is whether unidimensional income and multidimensional (non-income based) middle-class definitions identify the middle class consistently (both definitions identify the same individuals). Findings in Colombia by Castellani et al. (2015) demonstrate identification discrepancies between these approaches. The authors suggest the use of these inconsistencies to identify vulnerable individuals within the middle class (a lower middle class).

The main motivation behind this chapter is to analyse the middle class as a multivariate phenomenon instead of unidimensional. Parker (2013) and Wagley (2013) describe the middle class as a social group with a particular identity and characteristics beyond the economic domain. Castellani et al. (2015) argue that being in the “middle” of a single wellbeing domain’s distribution does not capture all other characteristics which could define the middle class.

Moreover, wellbeing has a multidimensional nature, so its analysis should consider more than just a single wellbeing domain (Maasoumi and Yalonetzky, 2013; Alkire and Foster, 2011). Then, this analysis considers the multidimensional nature of wellbeing, and introduces this notion in the definition of the middle class.

A multidimensional approach to middle class identification is more comprehensive than unidimensional income definitions. Besides, this approach could help to overcome identification biases in unidimensional definitions derived from economic shocks like unemployment. In terms of public policy, a multidimensional definition of the middle class could help to identify areas of vulnerability beyond income, such as social marginalisation or precarious labour conditions, in order to strengthen this social group.

The contributions of this chapter are in the area of the methodological definition of the middle class as a multidimensional phenomenon. Moreover, the use of the counting approach is popular for poverty identification (Castellani et al., 2005; Banerjeer and Duflo, 2008; Alkire et al., 2015: p 123), but no previous studies have used this approach to identify the middle class, making this method a novel contribution.

The first proposed methodological approach identifies that during 2012 in Mexico, 59.2% of individuals were middle class, 58.2% were deprived and 7.5% were affluent. Intersections between these classes allow identifying that 31.9% of middle-class individuals were lower middle class, and 10.7% upper middle class. The remaining 57.4% could be considered as mid-middle class.

When comparing the first proposed methodological approach to unidimensional income definitions, relative as well as absolute, no single definition has a perfect match, suggesting that the income middle class does not entirely identify the multidimensional dimension of the middle class and vice-versa. The absolute $10 - $50 USD PPP per day definition has the best matching to the counting approach at 34.0% of overlap (45,611 individuals).

Results on the second methodological approach show a high sensitivity of this method to changes in poverty thresholds. The highest incidence of middle class is found using $50 USD PPP per day (51.1%), and the lowest using $10 USD PPP (34.5%). Unfortunately, the characteristics of the multidimensional poverty definition in Mexico do not allow for comparisons between this method and the first methodological approach.

This chapter is divided in eight sections. The following section presents a literature review on the multidimensional nature of the middle class and its

situation in Latin America. Section three discusses the counting approach. Section four describes the first methodological approach. Section five explains the second methodological approach. Section six describes the Mexican socioeconomic context. Section seven presents the empirical analysis. Finally, the last section discusses the findings.

III.2 The multidimensional middle class, a literature review

The previous chapter discusses the identification of the middle class using a single wellbeing domain. This method allows comparisons of changes in the middle class between countries and over time. However, being in the “middle” of a single wellbeing domain's distribution does not capture all other characteristics which could define the middle class (Castellani et al., 2015).

Moreover, wellbeing has a multidimensional nature (Maasoumi and Yalonetzky, 2013). Therefore, a more comprehensive analysis of people’s wellbeing could, arguably should, be multidimensional in nature (Alkire and Foster, 2011). Key challenges in this multidimensional approach are: the identification of those wellbeing domains; and the identification of those in the “middle” based on some aggregation method that takes into account all the previously identified relevant domains. A more detailed discussion on these topics is developed in the following subsections.

III.2.1 The measurement of wellbeing

In classical economic theory, a definition of people’s wellbeing can be approached by the concept of permanent income. Defined as the pool of total lifetime resources at individuals’ disposal (Romer, 2012: pp 367), permanent income is limited by the impossibility of its direct observation (Friedman, 1957: pp 20).

Income is widely used as a wellbeing domain. However, it tends to be volatile over time due to its sensitivity to shocks (Stoyanova and Tonkin, 2018). Expenditure shows a more stable trend over time than income due to the persistence of habits or consumption inertia (Deaton and Muellbauer, 1990: pp 329). Expenditure can also provide a better approach to wellbeing than income (Meyer and Sullivan, 2011), and is a better measure of permanent income than current income (Birdsall, 2010). Indeed, people smooth

consumption expenditure in the face of income shocks through savings and credit (Zeller et al., 1997: p 12).

III.2.2 The middle class in Mexico and Latin America

The foundations of the modern analysis of social classes could be traced back to Marx and Veblen (Eisenhauer, 2008), who divided society into classes in terms of the control and possession of the means of production (Marx, 2015: p 56), and between those involved in productive activities and those self-excluded from them, referred to as the leisure class (Veblen, 2009: p 20).

As mentioned in the previous chapter, Veblen identified that aspirations in the lower class, the impecunious class, played an important role in its behaviour in the process of imitation of the leisure class called pecuniary emulation (Veblen, 2009: p 24). This process could follow consumption patterns, the ownership of certain assets and goods (Lopez, 2013), moral values (French, 2013) and educational achievements (Parker, 2013; Barros, 2013).

The definition of a set of characteristics which describe the middle class is not straightforward (De la Calle and Rubio, 2010). Some of these characteristics, nevertheless, are related to pecuniary emulation at different levels. Banerjee and Duflo (2008) identify, from a comprehensive survey of the literature on the middle class, a typical attitude in the middle class towards entrepreneurial and investment behaviours, as well as moral values, usually related to the leisure class. Social relations could also help to identify the middle class (Pike, 2013), although not necessarily from a pecuniary emulation process.

Wagley (2013) describes the middle class in Latin America as a group of costumers, who are used to having adequate accommodation, good dress, white-collar jobs in offices, classrooms, stores and other public places. Moreover, Waglye argues that the Latin American middle class desires the latest household appliances like television sets, fridges and washing machines, as well as automobiles, even though their economic reality does not match their desires.

Educational achievement has a relevant historical role in Latin America. In Mexico, during the 1970s, having a child attending the National Autonomous University (UNAM) was a sign of status related to the middle class (Parker, 2013). In Argentina, aspirations for professional studies led a whole generation to improve occupational conditions (labour intergenerational mobility) in the 1970s-80s (Barros, 2013). However, during the mid-1990s,

university studies could not guarantee a middle-income job due to an oversupply of professionals in Latin America (Parker, 2013).

The Organisation for Economic Co-operation and Development Latin American Outlook report of 2011 highlights, for the identification of the middle class, the relevance of employment characteristics and informality[20], the education level, attitudes toward entrepreneurship[21], home ownership and access to financial services[22] as important characteristics shared by the middle class (OECD, 2010).

Members of the middle class are also argued to share moral values (Banerjee, and Duflo, 2008). The capitalist spirit (Weber, 2001: p 10), a moral attitude towards investments based on patience and the ethics of perseverance as key to prosperity shared by the middle class, justifies its consolidation in society and the decline of the aristocratic leisure class (Doepek and Zilibotti, 2005).

However, this entrepreneurial perspective does not necessarily hold in the reality of the middle class, where individuals prefer a secure and stable well-paid job rather than the uncertainties of private entrepreneurship. Moreover, the entrepreneurial activities of the middle class are usually for self-subsistence and low-scale, which brings about few contributions to overall productivity (Banerjee and Duflo, 2008).

Del la Calle and Rubio (2010) analyse the Mexican middle class from a perspective of political preferences, following the 2006 national elections results. The authors argue that the moral values of the middle class are based on positive values from immigrants, honest informal workers, and hard workers who believe in the benefits of the meritocracy.

In Latin America, only 40% of people who self-identify as middle class would be classified as middle class using the 50% to 150% median relative definition. Among the remaining 60%, half are deprived population, and the other half belong in the affluent group (OECD, 2010). Eisenhauer (2008) finds that

---

20 Informality defined as working outside the formal sector, evading taxes or social security contributions.

21 Entrepreneurship is defined as the capacity to introduce innovative products and services to the market when established firms do not have the incentives.

22 Access to financial services helps a household in the process of capital accumulation through consumption credit and mortgages.

individuals in the USA self-identified as middle class have an income between 50% - 80% of the income median.

III.3 The counting approach

Following the arguments presented in the previous section, the middle class could be deemed a social group defined by multiple characteristics. Therefore, a multidimensional approach could be a more comprehensive way to identify this group rather than unidimensional approaches. This section presents a discussion of the counting approach to group identification, widely used in the measurement of poverty.

A significant challenge in multidimensional wellbeing analysis is the aggregation of multiple dimensions (Atkinson, 2003). Early poverty reports counted the number of different deprivations in distinct wellbeing domains related to a poverty condition for the UK (Townsend, 1979: p 251), followed by Belgium (Vranken, 2002) and Ireland (Layte et al., 2000). In these studies, a poverty score could be computed for each individual based on thresholds for each wellbeing domain.

Alkire and Foster developed the Multidimensional Poverty Index (MPI). This measure is based on counting the number of people's deprivations in different wellbeing domains giving them a particular weight, and then identifying the population as poor or not-poor by comparing their deprivation score against a poverty threshold (Alkire et al., 2015: pp 123). This method follows six steps:

1. Definition of the wellbeing domains.
2. Definition of a deprivation threshold for each wellbeing domain.
3. A binary identification of the deprivation condition for each wellbeing domain, taking the value of 1 when deprived and 0 otherwise.
4. Definition of a weight for each wellbeing domain.
5. Computation of a deprivation score, based on the weighted sum of deprivations.
6. Definition of a poverty threshold in the deprivation score.

A similar method is proposed in this chapter for the multidimensional identification of the middle class, with two plausible alternatives: a direct estimation, identifying directly the middle class based on its intrinsic characteristics; or an indirect estimation, based on the simultaneous and direct identification of the deprived and affluent classes, and then approaching the middle class as the residual difference between these two groups and the total population.

III.3.1 The counting approach in the middle-class identification

The counting approach proposes the class identification following four steps:

1. Selection of a set of characteristics within different wellbeing domains in order to define the deprived, middle and affluent classes.
2. The membership in each class for each characteristic is encoded into a categorical variable.
3. A class achievement index is computed by counting the number of codes in each categorical variable.
4. Finally, the individual classification into social classes is given by score thresholds on the class achievement index in a way similar to the MPI.

Three possible types of categorical variable could be found, which should fulfil the conditions of completeness (all individuals in the sample must be identified) and exclusiveness (no overlapping between categories):

1. Variables which only identify the deprived and the non-deprived in a particular wellbeing domain. These variables identify two conditions or statuses (deprivation and non-deprivation) within a given wellbeing domain. Then, for instance, $x_{ij} \in \{0,1\}$ stands for the value of variable x_j for individual i, where $x_{ij} = 0$ denotes deprivation and $x_{ij} = 1$ lack of deprivation.
   Examples of these variables could be access to basic services, chronic health conditions, lack of basic household assets, etcetera.

2. Variables which only differentiate the affluent from the non-affluent in a particular wellbeing domain. These variables identify two conditions (non-affluence categories and affluence) within a wellbeing domain. Then, for instance, $y_{ij} \in \{0,1\}$ stands for the value of variable y_j for individual i, where $y_{ij} = 0$ denotes non-affluence and $y_{ij} = 1$ affluence.
   Examples of these variables could be the consumption of conspicuous goods, ownership of financial stocks, etcetera.

3. Variables which differentiate between the deprived, the middle class, and the affluent in a particular wellbeing domain. Then, for instance, $z_{ij} \in \{1,2,3\}$ stands for the value of variable z_j for individual i, where $z_{ij} = 1$ denotes deprivation $z_{ij} = 2$ denotes middle-class, and $z_{ij} = 3$ denotes affluence.
   Examples of these variables could be the quality of the home building materials, years of education, etcetera.

Thus, an individual *i* can be categorised into a social class according to her "class score"[23], a weighted sum of the values in her categorical variables x_{ij}, y_{ij}, and z_{ij}. This class identification system leads to two main middle-class identification methods: a direct and indirect approach. Both methods are described in the following subsections.

III.3.2 Direct identification

The set $\{X, Y, Z\}$ can be further divided into necessary and non-necessary variables. Necessary variables are those exclusive to the identification of the middle class and cannot be substituted by any other variable. Non-necessary variables define conditions related to other social classes, aside from the middle class. However, non-necessary variables could help in identifying the middle class.

Let the number of necessary variables within each of the three groups described above be X_1, Y_1, Z_1 and the number of non-necessary variables be X_2, Y_2, Z_2, so for instance: $X_1 + X_2 = X$. Then the binary indicator ψ_n is equal to 1 if person n is middle-class. Otherwise, the indicator is equal to 0. Combining necessary and non-necessary variables, the latter involving a counting rule, the identification function is expressed as follows:

$$\psi_n^m = \prod_{j=1}^{X_1} \mathbb{I}(x_{nj} = 1) \prod_{j=1}^{Y_1} \mathbb{I}(y_{nj} = 0) \prod_{j=1}^{Z_1} \mathbb{I}(z_{nj} = 2) \left[\mathbb{I}\left(\sum_{j=1}^{X_2} \mathbb{I}(x_{nj} = 1) w_{xj} + \sum_{j=1}^{Y_2} \mathbb{I}(y_{nj} = 0) w_{yj} + \sum_{j=1}^{Z_2} \mathbb{I}(z_{nj} = 2) w_{zj} \geq k\right]\right. \quad \textbf{(III.1)}$$

where the weights (w) attached to the indicator functions ($\mathbb{I}(.) = 1$ if the content in parenthesis is true, otherwise $\mathbb{I}(.) = 0$) satisfy the following requirements: $w_{x1}, w_{x2}, \ldots, w_{zZ_2} > 0$ and $\sum_{j=1}^{X_2} w_{xj} + \sum_{j=1}^{Y_2} w_{yj} + \sum_{j=1}^{Z_2} w_{zj} = 1$. Also, $k \in [0,1]$ stands for the middle-class threshold. If $k = 1$, then effectively all variables are necessary. In the equation, m stands for the middle class.

Finally, the size of the middle class in the population is given by:

$$MC^D = \frac{1}{N} \sum_{n=1}^{N} \psi_n^m \quad \textbf{(III.2)}$$

where N is the total size of the population, and MC^D stands for the middle class as obtained from direct identification.

---

[23] This class score is calculated similarly as the deprivation score in the MPI.

III.3.3 Indirect identification

The indirect estimation method first identifies both the deprived and the affluent groups in a sample or population, and later infers the middle class as the residual expressed in the following equation:

$$MC^I = 1 - A - P \quad \textbf{(III.3)}$$

where *MC*^{I} stands for the middle class by indirect identification, *A* represents the affluent population, and *P* the deprived population.

The deprived can also be identified combining necessary and non-necessary variables, as in the middle-class case:

$$\psi_n^p = \prod_{j=1}^{X_1} \mathbb{I}(x_{nj} = 0) \prod_{j=1}^{Y_1} \mathbb{I}(y_{nj} = 0) \prod_{j=1}^{Z_1} \mathbb{I}(z_{nj} = 1) [\mathbb{I}(\sum_{j=1}^{X_2} \mathbb{I}(x_{nj} = 0) w_{xj} + \sum_{j=1}^{Z_2} \mathbb{I}(z_{nj} = 1) w_{zj} \geq k] \quad \textbf{(III.4)}$$

where p stands for deprived, and the weights and k do not need to be the same as in the case of equation (III.1). Then, the proportion of the deprived among the population is obtained as:

$$P = \frac{1}{N} \sum_{n=1}^{N} \psi_n^p \quad \textbf{(III.5)}$$

Likewise, the affluent can be identified:

$$\psi_n^a = \prod_{j=1}^{X_1} \mathbb{I}(x_{nj} = 1) \prod_{j=1}^{Y_1} \mathbb{I}(y_{nj} = 1) \prod_{j=1}^{Z_1} \mathbb{I}(z_{nj} = 3) [\mathbb{I}(\sum_{j=1}^{Y_2} \mathbb{I}(y_{nj} = 1) w_{yj} + \sum_{j=1}^{Z_2} \mathbb{I}(z_{nj} = 3) w_{zj} \geq k] \quad \textbf{(III.6)}$$

where a stands for the affluent, and the weights and k do not need to be the same as the deprived either as the middle class. Finally, the proportion of the affluent among the population is:

$$A = \frac{1}{N} \sum_{n=1}^{N} \psi_n^a \quad \textbf{(III.7)}$$

Where *A* stands for the affluent population.

III.4 First methodological approach: an application of the counting approach

This section presents an overview of the proposed method based on the counting approach to define and identify the middle class. A methodological contribution of this chapter in the middle-class identification is to use the counting approach for both direct and indirect identification.

III.4.1 Wellbeing dimensions

Five dimensions of wellbeing were considered for middle-class identification: education, health, dwelling characteristics, household assets and occupational characteristics. These dimensions were selected due to information availability in Mexico[24], besides following OECD (2010), the Colombian National Administrative Department of Statistics (DANE, 2018), Wagley (2013) and Parker (2013).

Then, these dimensions could be considered as the categorical variables in section III.3.1, and could be of the form of variables *x*, *y* or *z* depending on their nature. The definitions of these variables are detailed in the following subsections. Table III.1 shows a general overview of these variables and their wellbeing domains[25].

Table III.1 Dimensions and their wellbeing domains

Dimension	Wellbeing domains
Education	Technical or professional education
Health	*Seguro Popular* access or no access to health services Access to health services provided by the Social Security system or by private GP Private medical insurance
Dwelling characteristics	Overcrowding Lightbulbs
Household assets	Landline and mobile phones Cars, vans or pick-ups Hi-Fi, radios and music recorders TV VCR and DVD Computer and printer
Occupation characteristics	Low skilled occupations Skilled occupations Directive or top managerial occupations

24 A comprehensive discussion on this topic will be elaborated in section III.7.1 of this chapter.

25 Detailed information on these dimensions is found in the Appendix B.1.

III.4.2 Education

The definition of middle-class education is based on the social aspiration for higher education. According to Barros (2013) and Parker (2013), the aspiration for higher education is a distinctive feature of the Latin American middle class, working as an instrument for intergenerational mobility, access to better jobs and social status.

In Mexico, the educational system is divided into pre-school, for children in ages 3 to 6; primary school for children 6 to 12; secondary school for teenagers in ages 12 to 15; high school, technical education, or *bachiller* for ages 15 to 18; and professional or technical professional education (Ministry of Education, SEP, 2018). Following Barros (2013) and Parker (2013), the levels of education aspired by the middle class in the Latin American context are:

- **Normal education**. Studies to prepare professors and teachers at any level of basic and intermediate education. This education guarantees formal and stable jobs in the public education sector.
- **Technical education**. All technical and commercial studies which deliver a particular set of skills to perform highly specialised jobs.
- **Bachelor's degree**. Any Bachelor's degree, granted by a public or private university.
- **Postgraduate studies**. Any level of postgraduate studies like taught or research master's degree, PhD as well as medical specialisations.

The middle-class indicator of educational achievement measures whether at least one household member has, or is currently studying at, any of these aforementioned educational levels. Despite the equivalence between technical education and high school/*bachiller* concerning educational level, technical education stands out for its higher employability.

Once the educational condition for each individual is identified, a household-level[26] higher-education index (*HEI*) is calculated as follows:

$$HEI_h = \sum_{i=1}^{N_h} HE_{ih} \quad \textbf{(III.8)}$$

where HE_{ih} is equal to 1 if individual i in household h has completed higher education.

The absence of a household member with higher education identifies all household members as deprived in education. On the other hand, the middle

---

26 Household is defined as a group of individuals who live together in the same dwelling and share food.

class is defined as having at least one household member with higher education. The affluent is defined as having two or more members with fully finished higher education. The conditions for each category are:

1. **Deprived:** $HEI_h = 0$,
2. **Middle class:** $HEI_h > 0$.
3. **Affluent:** $HEI_h \geq 2$, only if the higher education studies are fully completed.

These categories are encoded in a *z*-type variable, where these categories are mutually exclusive and exhaustive.

III.4.3 Health

Access to health services is an essential condition of wellbeing (Nussbaum, 2011: pp 33), allowing people to have a proper functioning in the society as a mean to work (Robeyns, 2005), and its lack is related to a condition of social deprivation (DANE, 2018). In the case of catastrophic and chronic diseases, the lack of access to public health services could lead households to severe financial distress and vulnerability (CONEVAL, 2018: p 51).

Health services in Mexico are provided by public and private institutions, being the *Seguro Popular* (SP) a universal health safety net. Then, the definition of this dimension is based on the type of health institutions where people obtain their health care, and is divided into three categories:

A. **No access to state health services or affiliation to the SP.** No access to health services, or affiliation to the minimum, safety net, service.
B. **Affiliation to social security institutions or the frequent use of private general practitioners (GPs).** Affiliation to public health institutions through formal employment like the Mexican Institute for Social Security (IMSS), the Institute for Social Security and Services for State Workers (ISSSTE), services provided by the national oil company (PEMEX), and the army (SEDENA, SEMAR). This category also considers those people who visit a private GP, regardless of their health affiliations, because it shows purchasing power within the household.
C. **Private health insurance.** Affiliation to private health insurance. Sometimes this affiliation is offered to high-profile employees as work benefits.

To compute this dimension, a household average for each of these three categories (A, B and C) is calculated in the following way:

$$Ahealth_{kh} = \frac{\sum_{i=1}^{N_h} health_{ikh}}{N_h} \quad \textbf{(III.9)}$$

Where $Ahealth_{kh}$ stands for the household average usage of the *k* health category (A, B or C); $\sum_{i=1}^{N_h} health_{ikh}$ represents the sum of all *i* members which report usage of the *k* health category in the *h* household; and N_h is the household size.

To identify the class category to which individuals belong, a *z* type categorical variable is computed identifying the three, mutually exclusive and exhaustive, class categories. The conditions for each class category are:

1. **Deprived:** $Ahealth_{Bh} = 0$ and $Ahealth_{Ch} = 0$ or;
   $Ahealth_{Ah} > Ahealth_{Bh}$.
   A household is deprived in health when no members have access to health categories B or C, or when the household uses on average more A health services than B.
2. **Middle class:** $Ahealth_{Bh} \geq Ahealth_{Ah}$
   A household is middle class in health when, on average, uses more or equal B health services than A.
3. **Affluent:** $Ahealth_{Ch} > 0$
   A household is affluent in health when at least one household member has private insurance services C.

III.4.4 Dwelling Characteristics

Dwellings are the physical units where household members live. An essential characteristic of dwellings is that they can house one or more households. Therefore, all dwelling characteristics are assumed to be shared equally between all household members living in the same dwelling (CONEVAL, 2018: p 54).

This dimension considers two dwelling characteristics: an overcrowding index, following concerns on the effects of overcrowding on the household members (DANE, 2018; CONEVAL, 2018: p 55); and a light bulb index, an approximation of wealth considering the number of light bulbs in a household in a similar way as Doll et al. (2000) approach wealth based on the quantity of light countries irradiate.

Overcrowding is estimated as:

$$Overcrowding_h = \frac{N_h}{R_h} \quad \textbf{(III.10)}$$

Where $Overcrowding_h$ is the ratio between the total number of residents in the dwelling (N_h), divided by the total number of rooms used for sleeping (R_h)[27]. Then, the thresholds on this domain are:

1. **Overcrowded.** If $Overcrowding_h \geq 2.5$. [28]
2. **Normal.** If $2.5 > Overcrowding_h \geq 1$.
3. **Under-occupied.** If $1 > Overcrowding_h$.

Meanwhile, the light bulb index is calculated as:

$$Light_h = \frac{bulbs_h}{T_h} \quad \textbf{(III.11)}$$

Where $Light_h$ is an average of light bulbs in the household, $bulbs_h$ stands for the total number of light bulbs available in the household, and T_h gives the total number of rooms within the dwelling, regardless of whether they are used for sleeping or living. Then, a *z*-type categorical variable for this dimension is defined as:

1. **Low lightbulb usage.** If $0 \leq Light_h \leq 1.5$.
2. **Medium lightbulb usage.** If $1.5 < Light_h \leq 2.5$.
3. **Top lightbulb usage.** If$2.5 < Light_h$.

Then, using overcrowding and light bulb indices, households are classified in the overall *z* type variable of dwelling characteristics as follows:

1. **Deprived:** If the household is overcrowded or has a low-intensity light bulb usage.
2. **Middle class:** If the household has either a normal condition in the overcrowding index or a mid-intensity light bulb usage, while not being deprived in any of the two indicators.
3. **Affluent dwelling:** If the household is jointly under-occupied and has top-intensity light bulb usage.

While perhaps not immediately evident, the aforementioned three dwelling categories are both mutually exclusive and exhaustive.

---

[27] The multidimensional poverty measurement in Mexico has a similar overcrowding index; however, it considers all rooms in the household, whereas this estimation considers only rooms for sleeping.

[28] This is the same deprivation threshold used by CONEVAL (2018).

III.4.5 Household assets

The use of assets to infer household wealth is a common and intuitive approach (Filmer and Pritchett, 2001; Alikre et al., 2015), and used widely in the Latin American context (Attanasio and Szekely, 1999). This dimension is observed at household level, not at the dwelling level as the previous one. Per capita household asset estimation standardises the assets by the household size.

By definition, public goods are commodities for which the use by one agent does not preclude its use by other agents (Mas-Colell et al., 1995: p 359). All household members usually share household assets; however, their consumption could exclude other members' use (e.g. everybody can watch the TV, but not everybody can watch the programme of their choice at the same time; the use of the single household shower in the mornings).

Therefore, household assets are considered in this analysis as private goods, but their use is assumed to be shared evenly among all household members. This could be a strong assumption because, as shown by Deree et al. (2012), Latin America has a great gender inequality in asset ownership within households. However, this analysis assumes that the use of these assets is shared.

Thus, a per capita asset index is calculated as follows:

$$PCasset_{kh} = \frac{a_{kh}}{N_h} \quad \textbf{(III.12)}$$

Where a_{kh} is the total number of *k* type assets in household h. Then, each per capita value is split into three ordered categories as follows:

1. **Low ownership of asset *k*:** If $PCasset_{kh}$ < 0.2.
2. **Mid ownership of asset *k*:** If 0.2<= $PCasset_{kh}$ < 2.
3. **High ownership of asset *k*:** If 2 <= $PCasset_{kh}$.

Finally, the *z* type categorical variable for this mutually exclusive and exhaustive dimension is defined as:

1. **Deprived:** If the household has low ownership in at least half of its asset indices.
2. **Middle class:** If the household has mid ownership in at least half of its asset indices and has not been identified as a deprived household.
3. **Affluent:** If the household has high ownership in more than a half of its asset indices and has not been identified as a deprived or middle-class household.

III.4.6 Employment and occupation

Parker (2013), Crompton (2008: p 16), and Lopez (2013) highlight the association between labour activities in offices, professional white-collar activities, and moderate-to-high responsibility jobs with being part of the middle class. In this sense, a class definition based on labour conditions could be directly operationalised with the type of occupation.

In this sense, the occupational class identification is defined as follows:

1. The middle class would consider individuals in white collar administrative positions, skilled technicians, mid-range military and navy officials, and supervisors with administrative responsibilities in industrial activities. Higher education students in working age and retired population with a pension are also considered middle class.

2. The deprived population considers unskilled workers with intensive physical labour like machinery operators, workers in production lines, helpers in food preparation, peasants, low military and navy ranks, and labour intense related activities. This group also considers people involved in domestic work.

3. The affluent group is identified at the top of any private or public organisation like a nations' president, state governors, high-profile government officials, members of the parliament, CEOs, and managers or members of the board of directors. Landlords who profit from letting their properties are also considered as the affluent, as well as high-ranked officials like commanders, generals, Commodore and admirals.

Some manual specialised activities like artist or musicians could be misidentified as low qualified jobs. In order to control for this effect, specialised artisans, artists and musicians are considered as middle class. Unemployed and vulnerable non-economically-active population, like the retired without pension or the chronically ill, are not considered in this labour classification; however, if no other household member can be classified, then the whole household will share a deprived condition.

Assuming that all household members share occupational achievements, the definition of the *z* type categorical variable is:

1. **Deprived.** If the household has at least one deprived worker (or unemployed, or vulnerable non-economically-active member) and no other mid or affluent-level worker.
2. **Middle class.** If at least one household member is considered as a mid-level worker, student or retired with a pension.
3. **Affluent.** If at least one family member is an affluent-level worker.

This classification follows the possibility that household heads could have more desirable job-types compared to other household members; therefore the whole household would be in a better-off position despite other household members having a relatively lower occupational condition. In the opposite case, when the household head has a worse job compared to other household members, the benefits provided by the other better-off members will be transmitted to the whole household.

III.4.7 Aggregation of dimensions to identify the middle class

The class identification for each household is obtained combining all the five types of *z* categorical variables for each dimension described before. All these dimensions have the same weight, so no dimension is deemed more important than other. In this way, if the household has a shock (loss of health affiliation due to unemployment for instance), the other dimensions will help to maintain its relative class position.

The combination of these dimensions proceed as follows:

1. Construction of binary variables to identify each of the *z* variable categories for each dimension IDC_{hd}^{c}, where *IDC* stands for the binary variable for household *h* (1 if the household has the condition, 0 otherwise), in the *c* class (ρ deprived, m middle class, and a the affluent), for dimension *d*.
2. Compute the household score of class achievement (ω) for household *h*, by each *c* social class as follows:

$$\omega_h^c = \sum_{h=1}^{d} IDC_{hd}^{c} \quad \textbf{(III.13)}$$

Thus, each household will have a score of achievement for each class. A matrix of class achievement represents these scores, which looks as follows:

$$MA = \begin{bmatrix} \omega_1^p & \omega_1^m & \omega_1^a \\ \omega_2^p & \omega_2^m & \omega_2^a \\ \omega_h^p & \omega_h^m & \omega_h^a \end{bmatrix} \quad \textbf{(III.14)}$$

where *MA* stands for the matrix of class achievement for the whole sample, ω_h^p represents the overall deprivation score for household *h*; ω_h^m the overall middle-class score for household *h*; and ω_h^a the overall affluence score for household *h*.

To identify in which social class households belong, a direct approach would be to define thresholds on the ω_h^c scores as k^p for deprivation, k^m for the middle class, and k^a for the affluent. In this way, the middle class could be identified directly by k^m, or indirectly as in equation (III.3). Then, the class definitions are:

1. **Deprived**: ψ_n^ρ=1 if $\omega_n^\rho \geq k^p$, 0 otherwise;
2. **Middle**: ψ_n^m=1 if $\omega_n^m \geq k^m$, 0 otherwise;
3. **Affluent**: ψ_n^a=1 if $\omega_n^a \geq k^a$, 0 otherwise;

A perfect identification of these groups will lead to perfect independence and no intersections between sets p, m and a, fulfilling the conditions of completeness and exclusiveness. However, because the class identification depends on scores, and classes are commutative (all dimensions have the same weight, so it does not matter in which way they are counted) and associative (all classes are defined based on the same set of dimensions), then intersections between classes are a possible outcome.

III.5 Second methodological approach: a special case of the counting approach based on multidimensional poverty

This section presents a middle-class identification method inspired by Castellani et al. (2015), which combines a multidimensional poverty definition based on the counting approach to identify the deprived, plus a unidimensional middle-class income definition to identify the affluent.

Castellani et al. (2015) proposed the identification of vulnerable groups within the Colombian middle class by combining information on multidimensional poverty to relative definitions of the income middle class. Because this multidimensional poverty definition is not based on income, intersections between the multidimensional poor and the middle class are expected, which the authors defined as “vulnerable” middle class.

Inspired by Castellani et al. (2015), the proposed middle-class estimation based on multidimensional poverty follows the poverty definition in Mexico to identify the deprived population[29], and includes different relative and absolute middle-class income definitions to identify the affluent. Thus, the middle-class

[29] Unlike the Colombian case, the Mexican multidimensional poverty definition includes income as a wellbeing domain.

identification method based on multidimensional poverty could be considered as a special case of the indirect counting approach. The following subsections describe the Colombian and Mexican cases in detail, and subsection III.7.4 presents the empirical application for the Mexican case.

III.5.1 The Colombian case

In 2012, Colombia introduced a multidimensional poverty definition based on an adaptation of the Alkire-Foster MPI[30] (DANE, 2018). This approach consists of identifying 15 wellbeing domains grouped in five dimensions, giving each dimension the same weight. Table III.2 shows the different dimensions, their wellbeing domains, and their weights.

Table III.2 Multidimensional poverty measurement in Colombia, dimensions and wellbeing domains

Dimension	Wellbeing domain
Education conditions (0.2)	Educational achievement (0.1)
	Illiteracy (0.1)
Child and youth conditions (0.2)	School attendance (0.05)
	Educational backwardness (0.05)
	Lack of access to nursery services (0.05)
	Child labour (0.05)
Employment (0.2)	Long-term unemployment (0.1)
	Formal employment (0.1)
Health (0.2)	Access to health services (0.1)
	Social security coverage (0.1)
Household utilities and living conditions (0.2)	Access to clean water (0.4)
	Sewer connection (0.4)
	Adequate floor materials (0.4)
	Adequate walls (0.4)
	Overcrowding (0.4)

Source: Castellani et al. (2015), DANE (2018).
Note: particular weights for each dimension and wellbeing domain within brackets.

The poverty identification will aggregate the number of all the different weighted deprivations on these wellbeing domains in a score, and then set a *k* poverty threshold on this score domain. On the other hand, the affluent could be inferred by using any unidimensional middle-class definition described in section II.3 of the previous chapter.

30 Described in section III.3.1

A relevant characteristic is the absence of a monetary variable as a poverty dimension. Then, an overlap between the multidimensional deprived population and the income middle class is expected. Castellani et al. (2015) consider this overlap as the vulnerable condition of the middle class (e.g. an individual identified as income middle class but under multidimensional poverty conditions).

III.5.2 The Mexican case

In 2010, after a decade of academic and political debate, the Mexican Congress defined poverty as multidimensional, and commissions the National Council for the Evaluation of Social Development Policy (CONEVAL) to measure poverty every two years at national and state level, and every five years at the municipal level (CONEVAL, 2018: p 17).

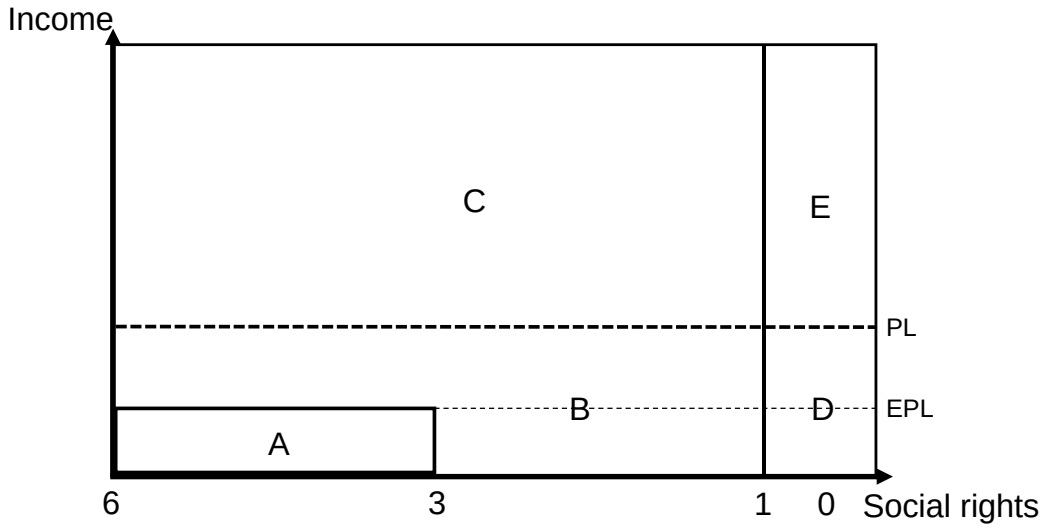
The main objective of this poverty definition is the evaluation and focalisation of social development programs. Two areas of deprivation are defined, one based on the lack of social rights, and a second one based on income (CONEVAL, 2018: p 19). Concerning social rights, a person is considered deprived if she does not fulfil the minimum requirements established in the law as follows:

1. **Education gap**, as not achieving the minimum standards on basic education.
2. **Access to health services**, as not been registered to health services.
3. **Access to social security**, as not covered by social security.
4. **Quality of spaces and dwelling**, as living in a dwelling with conditions below minimum standards of durability, protection against weather, and overcrowded.
5. **Access to basic services in the dwelling**, as living in a dwelling without utilities.
6. **Access to food**, as a subjective approximation to the households' experiences of lacking food.

Regarding income, a person is considered deprived if her monthly household per capita income is below a poverty line. The definition of the poverty line could be for extreme poverty or poverty (CONEVAL, 2012: p 14). The former represents the value of a food basket to guarantee the minimum of caloric and micronutrients consumption required by a person, whereas the latter represents the same basic basket plus the minimum requirements to cover expenditures on housing, clothing, education, health, transport, and recreation.

Figure III.1 shows a graphic representation of multidimensional poverty according to the official Mexican definition. Area A represents the population under extreme poverty as having at least three deprivations in social rights, and a monthly per capita income below the extreme poverty line (EPL). Area B (plus A) represents the population on poverty condition, with at least one deprivation in social rights, plus an income below the poverty line (PL). Area C represents the population with income over PL, but vulnerable in terms of social rights with at least one deprivation. Area D represents the population under income vulnerability, with no social rights deprivations, but income below PL. Finally, area E represents the population without social rights deprivations and income over the poverty line[31].

Figure III.1 Characterization of the population according to the Mexican official multidimensional poverty definition



Source: based on CONEVAL (2018: p 38).

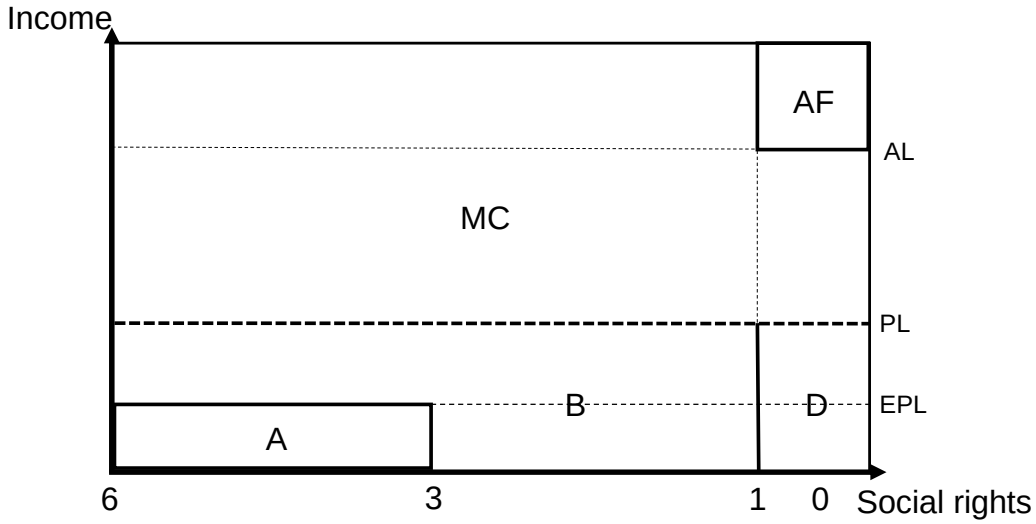
Following Castellani et al. (2015), an *affluent line* (AL) could be drawn to identify the affluent population based on the middle-class single wellbeing domain definitions discussed in section II.3 of the previous chapter. Then, a plausible identification of the middle class could be the population over poverty conditions but below the AL. However, the identification of the affluent could also be controlled by deprivation on social rights, not only by income.

A proposed identification of the middle class is drawn in Figure III.2, where the areas A, B and D remain the same as in figure II.1, the AF area represents the affluent population with income over the AL and no deprivation on social

31 Appendix B.2 reports values for the 2012 estimation.

rights, and the middle class (MC) is represented by the population with income over PL and below AL, or over AL but with at least one social right deprivation.

Figure III.2. Characterisation of the middle class and Mexican official multidimensional poverty



Source: Own elaboration based on CONEVAL (2018: p 38).

III.6 The Mexican context

This section presents a brief description of the social context in Mexico, starting with middle-class estimations based on unidimensional relative and absolute definitions, but using income and expenditure as wellbeing domains in order to gauge differences among them. Then, each of the five dimensions proposed for the counting approach estimation is discussed and analysed.

III.6.1 Unidimensional middle-class estimation in income and expenditure wellbeing domains

Despite its advantages as a wellbeing domain compared to income, a significant drawback of using expenditure is its reduced availability due to the challenges for its collection and inter-temporal consistency. In Mexico, the National Income and Expenditure Household survey (ENIGH) provides accurate information on both income and expenditure.

The unidimensional estimation of the middle class presented in this analysis follows the relative and absolute definitions discussed in section II.2.

Thresholds for relative definitions as percentages of the median are: 50% - 150% (Davis and Huston, 1992), 75% - 125% (Thurow, 1984), and 60% - 225% (Blackburn and Bloom, 1985). For absolute definitions, thresholds in daily USD PPP are: $2 - $13 (Ravallion, 2010), $10 - $50 (Lopez-Calva and Ortiz-Juarez, 2011), $2 - $10 (Banerjee and Duflo, 2008), and $12 - $50 (Milanovic and Yitzhaki, 2002). Table III.3 shows the estimations results for 2012. Monthly current household income and expenditure are calculated following equation (II.15) (in the previous chapter) and divided by 30.4 for per day estimations.

Table III.3 Middle-class estimation for relative and absolute definitions of income and expenditure, 2012

Definition	Income	Expenditure
Percentage of the median		
50% - 150%	46.7%	49.8%
75% - 125%	23.2%	25.2%
60% - 225%	53.4%	56.6%
USD PPP per day		
$2 - $13	59.8%	71.1%
$10 - $50	42.9%	28.4%
$2 - $10	48.7%	62.3%
$12 - $50	35.2%	21.9%
Total population	33,694	33,694

Source: own estimations using microdata from ENIGH, 2012.

Estimations based on expenditure show a higher incidence of middle class than those based on income, except for the absolute definitions with thresholds of $10 - $50 and $12 - $50 USD PPP. This effect could be explained by a wider income distribution, whereas expenditure has a flatter distribution[32]. If households have low levels of income, their expenditure should be lower compared to higher income households (Birdsall, 2010).

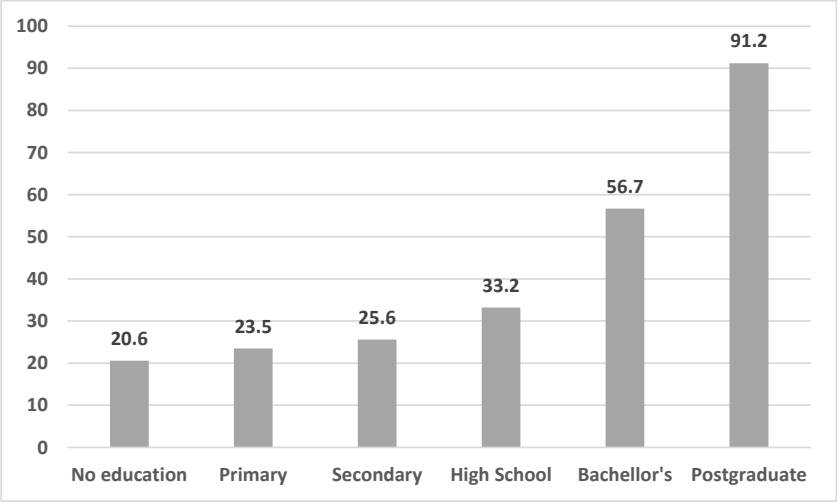
In the following subsections, the middle class is defined as income $10 - $50 USD PPP due to its extensive use among the literature as well as for being the official middle-class definition employed by the World Bank (Ferreira et al., 2013). Then, it is essential to keep in mind the monetary nature of the middle-class definition when analysing non-monetary dimensions of wellbeing.

32 Descriptive statistics, as well as density graphs, for both income and expenditure are presented in Appendix B.3.

III.6.2 Education

As discussed before, higher education is linked to better levels of income. Besides, it is related to a higher social status. Figure III.3 shows the average labour income per hour in Mexico for the 3rd quarter of 2011[33] by levels of education (no education, primary education, secondary, high school or equivalent, bachelors and postgraduate studies).

Figure III.3 Average income per hour by level of education, Mexican pesos, 3rd quarter 2011



Source: own estimations using microdata from ENOE, 2011.

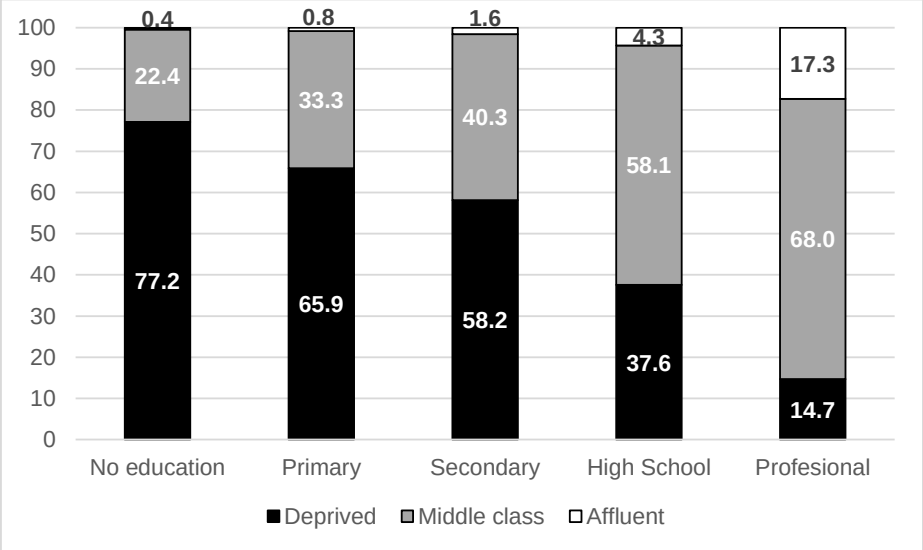
In this figure, an expected positive correlation between income and education level is found, reaching the top income at postgraduate studies ($91.2 Mexican pesos per hour). People with professional studies have on average an income 1.7 times higher than people with high school, 2.2 times more than secondary education, 2.4 times higher than primary education, and nearly 2.8 times more compared to people with no education.

Figure III.4 shows the percentage of class incidence by the level of education. A clear relationship between the higher level of education and better social condition is observed, particularly in the no education group, where most population are deprived, and the professional ones where most population are middle class or affluent.

---

33 Third quarter is observed as a period with low seasonality in the Mexican labour market (Rodriguez-Oreggia et al., 2013).

Figure III.4 Percentage of class incidence by education level, 2012



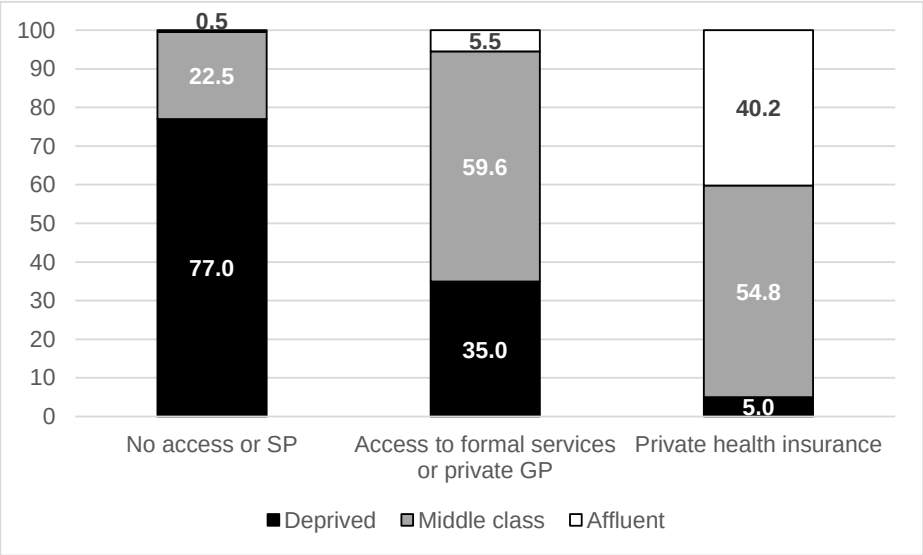
Source: own estimations using microdata from MCS-ENIGH, 2012.

III.6.3 Health access

In Mexico, social security (SS) provides access to public health services restricted to formal employment. In 2011, nearly 64.0% of all workers were not covered by social security (ENOE, 2018). Since 2003, a universal health system called *Seguro Popular* (SP) was created as a health safety net. However, SP services are restricted to a shortlist of common illnesses, excluding several chronic diseases and treatments which the SS covers.

Private health services are available to all population, but their access depends on the patients' budget. Private GPs are a popular option of private service, particularly in low-cost pharmacies which offer medical attention. Private health insurance is also available, but its demand is low. In 2015 only 24.8% of all adults had any type of insurance due to their high costs (ENIF, 2015: p 125). Figure III.5 shows the percentage of class incidence by type of access to health services in 2012.

Figure III.5 Percentage of class incidence by health service access, 2012



Source: own estimations using microdata from MCS-ENIGH, 2012.

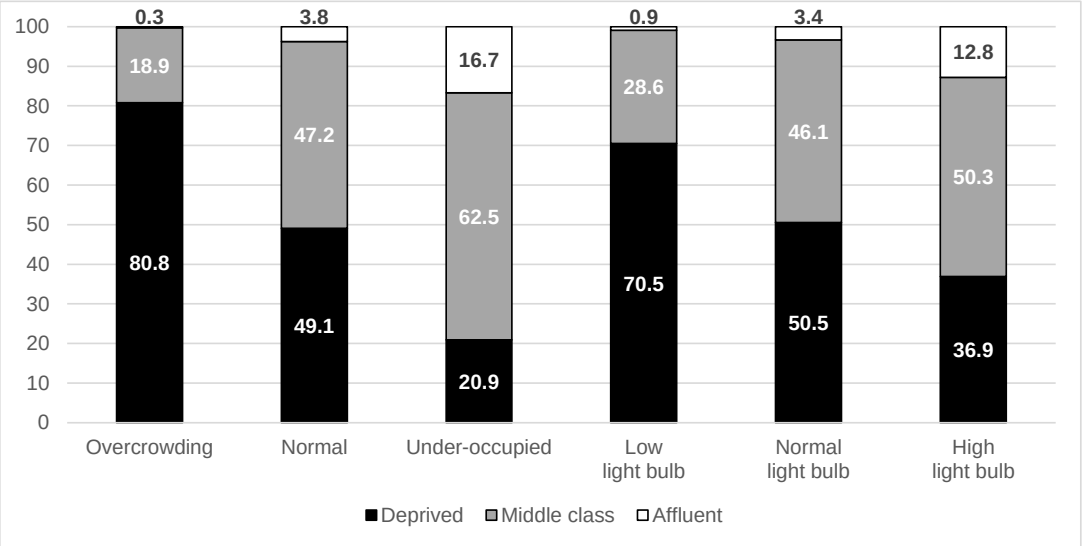
This graph shows high levels of exclusion on health services or use of SP by the deprived (77.0%), a broad dependence on SS and private GP services by the middle class (almost 60.0%), and an important coverage of private insurance among the affluent and the middle classes (40.2% and 54.8% respectively).

III.6.4 Dwelling characteristics

In 2012, more than 25.0% of the Mexican dwellings where located in rural areas (localities with less than 2,500 inhabitants), whereas 75.0% where in urban areas (MCS_ENIGH, 2012). Information on the characteristics of these dwellings is mainly restricted to the estimation of poverty and social marginalisation[34]. Considering this limitation, overcrowding and the number of light bulbs in the dwelling are considered as appropriate wellbeing domains. Figure III.6 shows the percentage of class incidence by these two domains as described in subsection III.4.4.

34 Meaning abundant *x*-type wellbeing domains, but not many *y*-type variables.

Figure III.6 Percentage of class incidence by overcrowding and light bulb indices, 2012



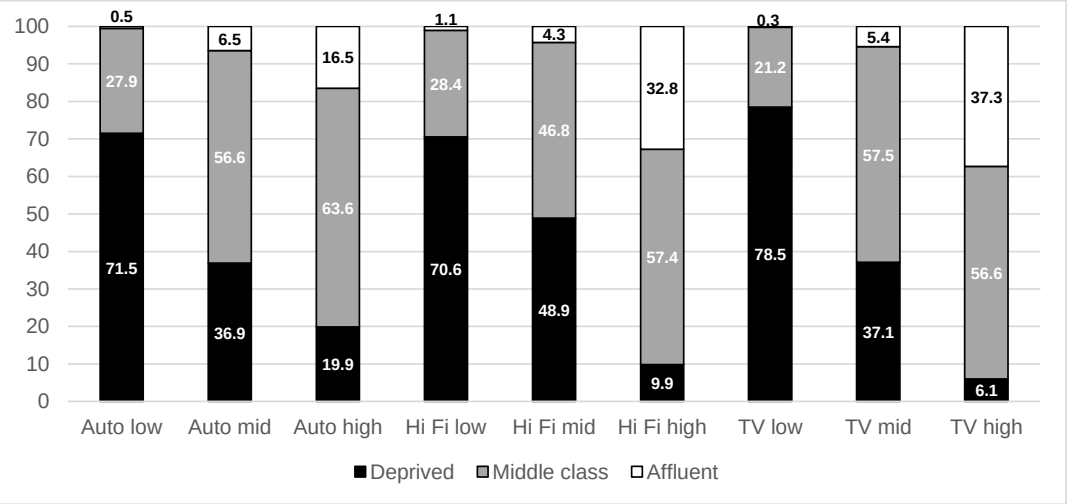
Source: own estimations using microdata from MCS-ENIGH, 2012.

A high level of overcrowding and low intensity of lightbulb usage is found in deprived dwellings (80.8%), while middle-class dwellings are largely under-occupied (62.5%) and have a high usage of lightbulbs (50.3%). The affluent class also has high levels of under-occupation (16.7%) and high lightbulb usage (12.8%).

III.6.5 Household Assets

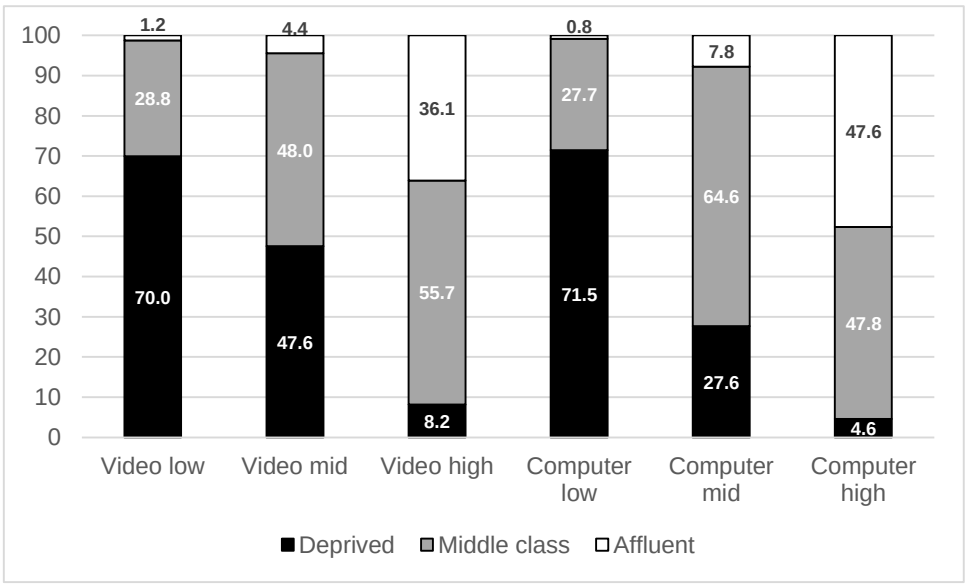
For the analysis of household assets, a household per capita asset index was estimated. The assets considered were vehicles, music systems, televisions, video players and computers, due to their relevance for the class definition, homogeneous usage countrywide, and information availability. The low, mid and high levels of ownership were computed as described in subsection III.4.5. Figure III.7 shows the percentage of class incidence by household ownership of automobiles, music systems and televisions, whereas Figure III.8 shows estimations for video players and computers.

Figure III.7 Percentage of class incidence by household asset (automobiles, music systems and televisions), 2012



Source: own estimations using microdata from MCS-ENIGH, 2012.

Figure III.8 Percentage of class incidence by household asset (video player and computers), 2012



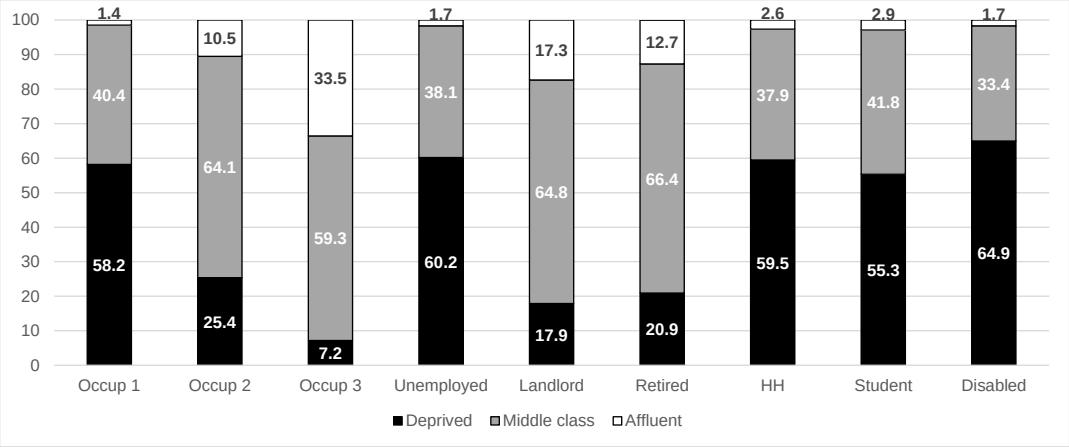
Source: own estimations using microdata from MCS-ENIGH, 2012.

These figures show a high incidence of deprived population in low ownership for all assets, particularly in TV (78.5%). The middle class has the most significant incidence in high ownership of automobiles (63.6%) and mid ownership of computers (64.6%). The affluent have almost the same incidence in high ownership of computers as the middle class (47.6%), and a high incidence in high TV and video player ownership (37.3% and 36.1% respectively).

III.6.6 Occupation

During the 3rd quarter of 2011, 60.7% of the Mexican population was economically active, and the national unemployment rate was 5.5% (ENOE, 2018). Figure III.9 presents the percentage of class incidence by type of occupation, considering Occup1 as low skilled-blue collar workers, Occup2 as mid-level white collar workers, Occup3 as high-profile workers[35], unemployed population, landlords, retired who receive a pension, housekeeping workers (HH), students and the disabled to work.

Figure III.9 Percentage of class incidence by occupation, 2012



Source: own estimations using microdata from MCS-ENIGH, 2012.

The deprived population shows high incidences in vulnerable labour activities like housekeeping and unskilled blue-collar jobs (58.2% and 59.5% respectively). This social class also has the highest incidences in unemployment, students and the disable (60.2%, 55.3% and 64.9% respectively). The middle class has a strong presence in white-collar and high-level jobs, landlords and the retired who receive a pension (64.1%, 59.3%, 64.8% and 66.4% respectively). Finally, the affluent have a high incidence in high-profile jobs and landlord (33.5% and 17.3% respectively).

III.7 Empirical analysis

This section develops the empirical identification of the middle class based on the proposed counting approach, as well as following the multidimensional poverty approach, for Mexico in 2012. This year was selected in order to

35 CEOs, directors, high-level public servants, etcetera.

identify the middle class which would work as a reference group for its identification in the samples used in the empirical analysis of chapter four.

III.7.1 Data

The Mexican National Institute for Geography and Statistics (INEGI) publishes the National Household Income and Expenditure Survey (ENIGH) every two years since 1992. This survey captures detailed information on sociodemographic characteristics of the household members, dwelling characteristics, household income and expenditure among other information, and is represented at national, urban and rural levels[36].

However, this survey does not capture all the information required to calculate the multidimensional definition of poverty discussed in section III.5.2. Then, a new survey, the Module of Socioeconomic Conditions of the ENIGH (MCS-ENIGH), was published biannually since 2008. The main difference between ENIGH and MCS-ENIGH is the absence of information on household expenditure in the latter.

Due to these characteristics, the 2012 sample of the MCS-ENIGH was selected for the counting approach estimation of the middle class. Moreover, being this sample the data source for the multidimensional poverty estimation, the indirect identification of the middle class described in section III.5.2 is computed straightforwardly.

III.7.2 First methodological approach, estimation results

This estimation follows the method described in section III.4, using the dimensions in Table III.1, and households as units of analysis. Asset ownership considered the following items: Cars, vans and pick-ups; Hi-Fi music components, radios and music recorders; TV; VCR, DVD and *Blu-ray* equipment; and computers and printers.

The lack of household landline and mobile phones is used as an *x*-type variable for the overall dimension. In the case of vehicles, an extra *y*-type control was introduced in the form of car ownership, whereby having two or more vehicles with a maximum age of two years is deemed symptomatic of affluence. For all other assets, only per capita values were computed.

---

36 In some particular years, information at the state level is also provided.

The selection of these assets depends on data availability. As mentioned before, this survey is mainly aimed at identifying poverty; thus no high-end assets were found as *y*-type variables. Assets like fridge, washing machine, air-conditioning systems and heating lead to a bias in the identification of the deprived population in rural areas and regions with extreme weather.

Table III.4 shows the percentage distribution of classes by each dimension (frequencies of the *z* type variables for each dimension). Education is the dimension with the lowest achievement, where 72.7% of households are deprived, and only 26.0% are in the middle-class condition. Health, on the other hand, has the highest middle-class achievement (56.8%) and lowest deprivation (41.0%). Dwelling characteristics and household assets also have high levels of deprivation, where 57.4% and nearly 59.6% of households are deprived respectively. Finally, occupation is the dimension with the most substantial affluent achievement, 3.2% of all households[37].

Table III.4. Social class achievement by dimensions, percentages

Classification	*z* Education	z Health	z Dwelling	z HH assets	z Occupation
Deprived	72.7	41.0	57.4	59.6	50.3
Middle class	26.0	56.8	42.2	38.9	46.5
Affluent	1.3	2.2	0.4	1.5	3.2
Total households	**56,424**	**56,424**	**56,424**	**56,424**	**56,424**

Source: own estimations using microdata from MCS-ENIGH, 2012.

Once identified the degrees of social class achievement by dimension, binary variables were created to identify each household level of achievement in each dimension. Then, these binary variables are used to calculate a matrix of achievement for each social class following equation (III.14). Table III.5 shows the percentage distribution of these matrices of achievement in scores.

Table III.5 Percentage distribution of the achievement scores by social class

Score values	Deprivation scores	Middle class scores	Affluent scores
0	11.9	19.2	92.5
1	13.3	21.6	6.4
2	15.8	19.6	1.0
3	18.7	17.1	0.2
4	21.3	14.1	0.0
5	19.0	8.3	0.0
Total households	**56,424**	**56,424**	**56,424**

Source: own estimations using microdata from MCS-ENIGH, 2012.

The maximum score of class achievement a household can obtain is 5, whereas the minimum score is 0. The distribution of deprivation scores has

[37] Descriptive statistics for the wellbeing domains and a correlation matrix between dimensions are shown in Appendix B.4.

the highest concentration of scores 5, meaning 19.0% of households are deprived in all dimensions. Thus, 8.3% of households are middle class in all dimensions, and no household in the sample is affluent in all dimensions.

Most households (72.7%), however, have no unique identification in a single social class but present a combination of different scores for each class. To classify them within overall classes, *k* thresholds are defined as:

- k^p**:** for the deprived, at deprived score >= 3.
- k^m**:** for the middle class, at middle-class score >= 2.
- k^a**:** for the affluent, at affluent score >= 1.

After the classification using the *k* thresholds, the three social classes are finally integrated. Table III.6 shows their incidences in percentages.

Table III.6 Percentage of incidence by social class

Statistic	Deprived	Middle class	Affluent
Mean	58.2	59.2	7.5
SE	0.21	0.21	0.11
Total households	**56,424**	**56,424**	**56,424**

Source: own estimations using microdata from MCS-ENIGH, 2012.
Note: SE stands for standard error on the mean.

The identification criteria fulfil the condition of completeness for all classes, albeit not exclusiveness due to a natural intersection among classes as stated in subsection III.4.7. Table III.7 shows frequency cross-tabulates between classes.

Table III.7 Cross frequencies between social classes

Middle class and deprived

	No deprived	Deprived	Total
No middle	470	22,574	23,044
Middle class	23,095	10,285	33,380
Total	23,565	32,859	56,424

Middle class and affluent

	No affluent	Affluent	Total
No middle	22,574	470	23,044
Middle class	29,638	3,742	33,380
Total	52,212	4,212	56,424

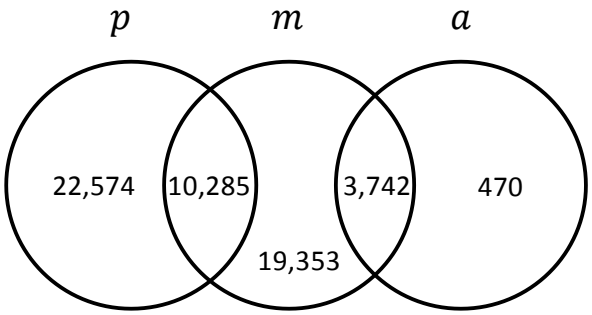
Deprived and affluent

	No affluent	Affluent	Total
No deprived	19,353	4,212	23,565
Deprived	32,859	0	32,859
Total	52,212	4,212	56,424

Source: own estimations using microdata from MCS-ENIGH, 2012.

Considering sample *N*=56,424 households, intersections are observed between the deprived and the middle class $NP(p \cap m) = 10{,}285$, and between the middle class and the affluent $NP(m \cap a) = 3{,}742$. No intersection is observed between the deprived and the affluent. Households in the set $p \cap m$ could be considered as a lower middle class or, following Castellani et al. (2015), a middle class in vulnerable condition. Households in $m \cap a$ could be considered as an upper middle class. Figure III.10 presents these results in a Venn diagram.

Figure III.10 Intersections between social classes



Finally, Table III.8 reports the direct and indirect estimations of the middle class. For the first approach, the middle-class incidence is the same as in Table III.6, 59.2% (33,380 households). For the second approach, the middle class is computed following equation (III.3), giving an incidence of 34.3% (19,353 households).

Table III.8 Percentage middle class incidence, direct and indirect approaches

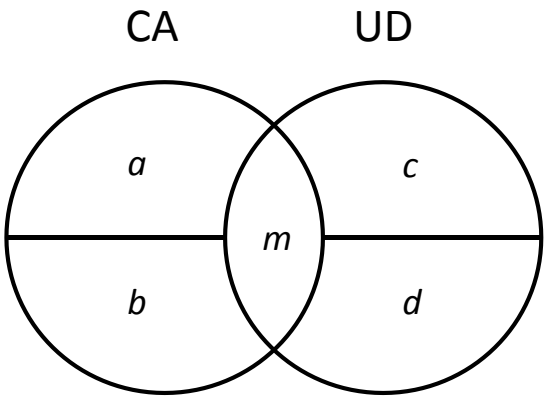
Statistic	Direct	Indirect
Mean	59.2	34.3
SE	0.21	0.20
Total households	**56,424**	**56,424**

Source: own estimations using microdata from MCS-ENIGH, 2012.
Note: SE stands for standard error on the mean.

III.7.3 Intersections between estimation approaches

As shown by Castellani et al. (2015), intersections between middle-class estimations based on income and non-monetary wellbeing domains could help to identify groups in condition of vulnerability. A proposed way to analyse these intersections is presented in Figure III.11.

Figure III.11 Analysis of intersections between different middle-class definitions



CA stands for the set of observations identified as middle class by the counting approach; UD for a set of observations defined as middle class using a unidimensional definition, and the subsets *m, a, b, c, d* which are interpreted as:

- *m*: the intersection of CA and UD (CA ∩ UD)
- *a*: the counting approach upper middle class (income > affluence line)
- *b*: the counting approach lower middle class (income < deprivation line)
- *c*: the unidimensional definition affluent (affluent score >= 1)
- *d*: the unidimensional definition deprived (deprivation score >= 3)

Table III.9 presents the estimated values of these areas in percentages for the counting approach and the income definitions described in subsection III.6.1.

Table III.9 Analysis of intersections between counting and income middle-class definitions, percentages

	Percentage of the median			USD PPP per day			
Condition	**50% - 150%**	**75% - 125%**	**60% - 225%**	**$2 - $13**	**$10 - $50**	**$2 - $10**	**$12 - $50**
Matching	17.5	11.2	26.6	14.5	34.3	9.7	30.8
Counting approach							
Upper middle class	36.2	50.3	22.7	35.0	6.6	45.4	6.9
Lower middle class	5.8	16.4	8.9	0.7	28.6	0.7	38.1
Unidimensional income definition							
Affluent	1.6	1.0	3.0	1.2	7.0	0.8	6.8
Deprived	39.0	21.2	38.8	48.5	23.4	43.4	17.5
Total households	**44,708**	**38,341**	**43,684**	**51,431**	**38,194**	**49,447**	**36,735**

Source: own estimations using microdata from MCS-ENIGH, 2012.

These results show no perfect identification of the middle class between both methods. The $10 - $50 USD PPP per day definition has the best matching to the counting approach (34.3%, 13,119 households). This estimation also adjusts better for the upper middle class (6.6%), although overestimates the lower middle class (28.6%) and the affluent (7.0%). On the other hand, the poorest matching is given by the $2 - $10 USD PPP definition (9.7%).

III.7.4 Second methodological approach, estimation results

The estimation of the middle class based on the official poverty measurement follows the methodology discussed in subsection III.5.2. Table III.10 presents the estimation results for poverty and extreme poverty as deprivation thresholds, meaning that for the later the middle class will be defined as the area MC plus B and D from Figure III.2. The affluent threshold is based on absolute and relative income definitions.

Table III.10 Middle-class estimation for the official poverty approach, using different poverty and affluent thresholds

Definition	Poverty	Extreme poverty
Percentiles		
> p(80)	41.6	78.2
> p(60)	34.7	71.2
> p(90)	46.6	83.2
Percentage of the median		
150%	37.1	73.6
125%	34.8	71.3
225%	42.7	79.2
USD PPP per day		
$13	37.9	74.4
$50	51.1	87.7
$10	34.5	71.0
Total population	**212,674**	**212,674**

Source: own estimations using microdata from MCS-ENIGH, 2012.

The definition column refers to the affluence thresholds based on the absolute and relative definitions of section III.6.1. As expected, the middle class estimated with extreme poverty is larger than using poverty due to the addition of the areas B and D. Estimations are also sensitive to affluence thresholds, as higher the threshold is defined as larger the middle class is. In general, estimations are consistent with unidimensional estimations in table III.3.

Finally, these estimations are not comparable to the counting approach results because the unit of analysis in multidimensional poverty are individuals, not households, having the possibility of individuals in different poverty conditions (one poor and another one no-poor) living in the same household. In the counting approach, class conditions are assumed to be shared between all household members.

III.8 Conclusion

This chapter analyses the middle class as a multidimensional phenomenon, and proposes a novel identification method based on the counting approach, popular in the field of poverty measurement but never used before for middle class identification. As discussed in this chapter, the multidimensional definition of the middle class is more comprehensive than unidimensional approaches. Moreover, giving all dimensions the same weight, economic shocks could have no effect on the middle class identification.

Two methods for class identification were proposed, following the counting approach, to identify the middle, the deprived and the affluent classes. The empirical analysis is based on a sample of the Mexican survey MCS-ENIGH for the year 2012.

The first method is based solely on the counting approach and has two ways to approach the middle class: a direct identification of the middle class based exclusively on its characteristics, or an indirect identification as the residual population group between the deprived and the affluent.

The second method, inspired by Castellani et al. (2015), can be considered as a special case of the counting approach. This method identifies the middle class as the population above the multidimensional poverty condition and below the top threshold of a unidimensional income definition of the middle class.

In the case of the first method, the direct estimation of the middle class identifies 59.2% of individuals in the sample as middle class, whereas 58.2% and 7.5% are identified as deprived and affluent respectively. An expected intersection between the groups, due to the aggregation of dimensions, is observed particularly between the middle class and the deprived and affluent groups. No intersection was found between the deprived and the affluent.

These intersections can be interpreted as a lower and upper middle class when intersecting the middle class with the deprived and the affluent respectively. Thus, 31.9% of the middle class can be identified as lower middle class, and 10.7% as the upper middle class. The remaining 57.4% can be considered as mid-middle class.

When comparing the counting method to relative and absolute definitions of the income middle class, estimations show no consistency between them. This inconsistency is similar to that of the Colombian case found by Castellani

et al. (2015). Then, the income middle class does not identify the multidimensional dimension of the middle class completely.

From all income-based middle-class definitions tested, the absolute $10 - $50 USD PPP per day has the best matching to the counting approach at 34.0% (representing 45,611 individuals). This estimation identifies well the upper middle class. However, this definition overestimates the lower middle class and the affluent. On the other hand, the poorest matching is given by the $2 - $10 USD PPP definition, suggesting that this definition does not reflect the real conditions of the middle class in a middle-income country, as suggested by Birdsall (2010).

The second method, based on poverty as an indirect definition and upper-income middle-class thresholds, shows a high sensitivity to changes in poverty thresholds. The highest incidence is found using $50 USD PPP per day (51.1%), and the lowest using $10 USD PPP (34.5%). Unfortunately, the characteristics of the multidimensional poverty definition in Mexico do not allow for comparisons between this method and the first method discussed previously.

One of the main challenges faced when applying the counting approach is the absence of enough and complete information for the definition of *y*-type categorical variables. This issue comes as a result of political prioritisation of poverty alleviation. Then, surveys are designed to collect information about vulnerable social groups rather than on the top-end of society.

Future research could consider a regional analysis of the middle class. As observed, the distribution of certain household assets does not follow a normal homogeneous distribution across the country; instead, assets are segmented according to cultural and environmental conditions. This correlation could be used as selection criteria for a regional middle class.

Another topic for future research is the use of the international MPI poverty dataset, which samples multidimensional poverty in a wide range of countries, to identify the deprived class. In this way, an international comparison of the middle class, with consistent class definitions, could be achieved.

Chapter IV
No-fault divorce and female labour market participation: evidence from the Mexican middle class[38]

IV.1 Introduction

This chapter assesses the impact of the introduction of no-fault divorce in Mexico on the labour market participation of women belonging to the middle class, focusing on three dimensions: overall labour participation, informal sector participation and the number of hours worked per week.

The middle class in this chapter is identified following a unidimensional absolute definition, using a predicted labour income as a wellbeing domain which considers the multidimensional nature of the middle class as discussed in the previous chapter. The middle class is particularly important for both economic and institutional development. Therefore, changes in the behaviour of the middle class could have repercussions in the whole society[39].

Previous studies have analysed these effects in the context of developed countries, but no studies exist for developing countries. Moreover, no previous studies have analysed these effects on women from the middle class, making the analysis on this chapter pioneer in the field.

The main research question in this chapter is whether the introduction of no-fault divorce has an effect on women's labour market participation, and its implications on participation in the informal labour market[40] as well as the hours worked per week.

The informal labour sector is often associated with precarious working conditions and vulnerability due to the absence of social security, job stability and lawful guarantees (Rodriguez-Oreggia et al., 2013). It is also a type of sub-employment (ILO, 2014). In Mexico, more than 61.0% of the total working population and nearly 60.3% of the total female working population

---

38 A different version of this chapter, without concentring just on the middle class, has been presented at the Asian Meeting of the Econometric Society 2018, held in Seoul (21-23 July), and went through refereeing process.

39 Wagley (2013) considers the Latin American middle class as a progressive group which, facing live in an industrialised urban society, constantly struggles against the conservative values of the establishment. This struggle guides the middle class towards "modern" behaviours, which could inspire other social groups.

40 Informal labour market lacks social security coverage following Mexico's Labour Law.

participated in the informal sector during 2011 (ENOE, 2011). Hence, the relevance of focusing on the labour sector which represents the largest share of working population, and where entry jobs are usually found.

The main motivation behind this chapter is to contribute in the analysis of the middle class in its labour dimension, to provide evidence for developing countries, as well as to contribute in the field of quasi-experimental analysis with an alternative empirical estimation strategy to traditional estimations when panel data is not available.

The introduction of no-fault divorce in the USA had different outcomes on women's conditions, depending on the institutional background. Jacob (1989) finds little or no effect of no-fault divorce introduction on women's financial situation. Genadek et al. (2007) identify in the USA a positive significant effect on labour participation of married women with children compared to married women without children. On the other hand, Parkman (1992) shows that no-fault divorce legislation with a narrow definition of property rights and limited alimony has a positive significant influence on women's labour participation and their investment in human capital in some USA States.

Meanwhile, developing countries are characterised by the absence of social welfare systems with universal outreach, high levels of labour informality, and weak judicial institutions. Therefore, the effects of introducing no-fault divorce on female labour supply might be different in developing countries vis-a-vis developed countries (as well as within each group of countries). Changes in divorce legislation could have different effects in developing countries. For instance in Chile, the introduction of divorce in 2004 increased the women's probability of having their first child by 61.0% (Gallegos and Ondrich, 2017). However, no previous studies have focused on the particular effect of no-fault divorce introduction on women's labour market participation in developing countries.

No-fault divorce was introduced in Mexico in 2008 as an attempt to make the process of marital dissolution accessible and easier for most of the population, thereby seeking to empower women through the promotion of bargaining power within households, thus enhancing human rights, freedoms, and their capabilities, following recommendations from the National Human Rights Commission (CNDH), and the National Supreme Court of Justice (SCJN) (CNDH, 2013; SCJN, 2012).

This divorce legislation was introduced gradually due to Mexico's federal organisation[41], starting with the Federal District (DF) in 2008, then, followed by six other States between late 2011 and 2013, until becoming universal in 2015. Hence, this setting is amenable to a quasi-experimental difference-in-difference (DID) econometric analysis following Genadek et al. (2007), where individuals and their characteristics do not influence the law's implementation, i.e. the event is completely exogenous to them.

This DID estimation strategy defines women living in the DF as 'treatment' observations, whereas women living anywhere else are members of a 'control' group, for two periods of time: 2005 (i.e. three years before the law's introduction) and in 2011 (i.e. three years after the law's introduction)[42]. Among 'treated' middle-class women, an average increase in labour market participation of around 3.5 percent points was found as an effect of the no-fault divorce introduction, alongside a reduction in informal labour market participation of 3.1 percent points, and an increase in working hours of 4.0%.

This chapter is divided in six sections. The following section gives a general context on divorce, labour market and social classes in Mexico. Section 3 presents the literature review. Section 4 discusses the methodology. Section 5 describes the data used in the empirical analysis and presents the estimation results. Finally, section 6 discusses the findings of this chapter.

IV.2 Mexican context

This section presents a brief overview of the Mexican context. The first subsection describes the legal framework behind the no-fault divorce in Mexico as well as its trends in recent years. The second subsection analyses trends on labour market indices like net labour participation and unemployment rates. The final subsection shows a socio-demographic

41 Mexico has 31 Federal States and a Federal District (Distrito Federal in Spanish), which is the political capital of the country and is regulated by its own laws. Considering its population, geography and economic characteristics, the Federal District is considered as one State in this analysis, leading to 32 States.

42 The selection of this timeframe follows reasons of data availability and comparability. In similar studies, the timeframe follows decades before and after the law innovation using census data. However, the Mexican census data required for a similar analysis will not be available until 2021. On the other hand, a comparable series of labour surveys started only in 2005, making it impossible to analyse previous years.

characterisation of women compared to men, and of middle-class women compared to deprived and affluent women.

IV.2.1 Divorce in Mexico

Divorce in the Mexican legal framework exists since 1917, defined by the Federal Civil Law (FCL) as the legal procedure to dissolve marital links between spouses, enabling them to contract another marriage (FCL, 2013). This Law establishes twenty causes of divorce from which at least one should be proven. Nevertheless, the thirty-two State legislatures adapt the Federal Law, following local requirements based on idiosyncrasy and culture, thereby adding more causes to this list.

In practice, proving these accusations is difficult and linked to non-desirable socio-emotive implications like personal denigration, exposition of private details, and harmful confrontations (SCJN, 2012). Moreover, divorce procedures are expensive and time-consuming: the most straightforward divorce agreement could cost nearly three times a monthly minimum wage[43], and last between one to two years (Aguillon-Leon et al., 2010). In Latin America, divorce appears as a penalty or exception for people, limiting their freedoms (Marco, 2009).

No-fault divorce, known in Mexico as '*divorcio exprés*', is a type of divorce which does not require proving any cause, can be demanded unilaterally by any of the spouses, and is usually resolved in less than a year at a low cost (SCJN, 2012), making this procedure accessible to all the population, particularly for economically deprived women.

Initially established in some USA states[44] and European countries between 1960 and 1987, no-fault divorce was introduced in Mexico in the DF on 2008. The State of Hidalgo introduced it at the end of 2011, followed by the State of Guerrero in 2012, and then by the States of Coahuila, Mexico State, Sinaloa and Yucatan during 2013. Finally, it became universal in 2015.

In the USA, a short-term exposure of 3 years to no-fault divorce increased the local divorce rate (Nakonezny et al., 1995). However, long-term exposure has no effects on divorce rates (Allen, 1998). Moreover, prolonged exposure to

---

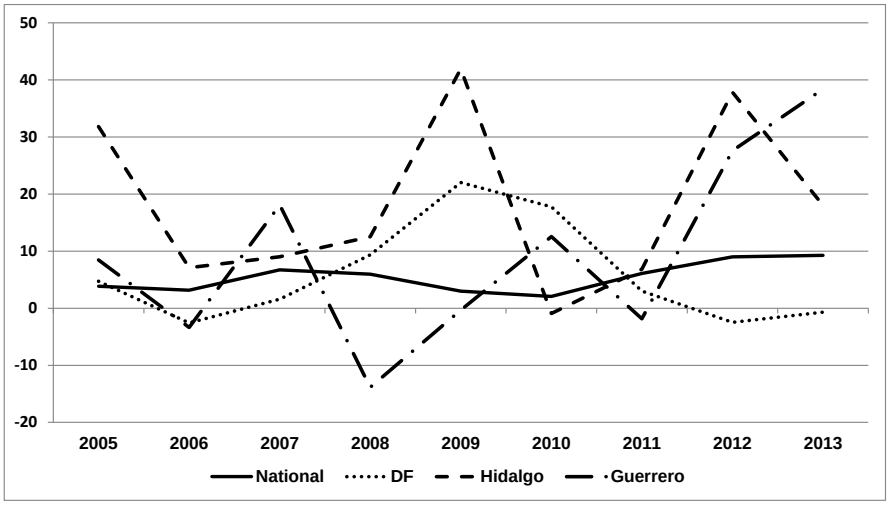
43 Considering an average minimum wage of $1,595.70 pesos per month in 2009, published by the National Minimum Salary Commission (CONASAMI, 2017).

44 Becoming universal in the USA in 1987 (Hill-Kay, 1987).

no-fault divorce led people to get married and diminished the chances of getting divorced (Johnson and Mazingo, 2000). In Western European countries, a positive relationship is found between no-fault divorce implementation and the divorce rate (Kneip and Bauer, 2009).

Figure IV.1 shows the growth rate of divorce at national level and for the States of Hidalgo, Guerrero and the DF, which are the areas with data on divorce available for the period when no-fault divorce was introduced. DF had an increase on the divorce rate in 2009 and 2010 when no-fault divorce was introduced, but the rate decreased to a similar level before the implementation. Hidalgo has a peak in 2009, probably influenced by the DF, returning to its original levels in the following years, then, peaking up in 2012 after their legislation changed, and finally decreasing. In Guerrero, a similar trend is observed, an increase in divorce rates when the law was introduced.

Figure IV.1 Annual growth rate of divorce, National, DF, Hidalgo and Guerrero, 2005 to 2013



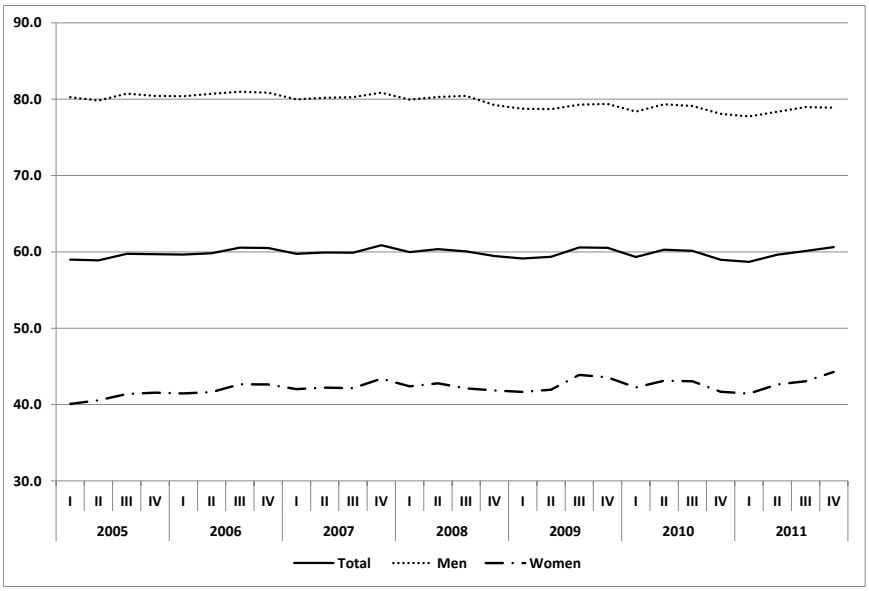
Source: Information data bank, INEGIa, 2013.

Then, despite the widespread belief from conservative groups in the Mexican society whereby introducing no-fault divorce would increase divorce rates by facilitating this process, trends on the annual growth of divorce for States where no-fault divorce was introduced increased when the law was implemented, but then decreased over time to levels similar to those observed before the implementation, suggesting an increase in divorce only in the short run, but no permanent effect. These results are consistent with the dynamics found in the USA by Nakonezny et al. (1995), Johnson and Mazingo (2000), and Allen (1998). An analysis of divorce rates using longer data series could help to clarify this relationship.

IV.2.2 The labour market

Regarding the Mexican labour market, women participate less than men. Figure IV.2 shows the quarterly total net labour participation, for men and women, from 2005 to 2011. Nearly 80.0% of men of working age (15-64 years) were participating in the labour market, while around 41.0% of women of the same working age were working, and 60.0% of the overall population in working age were working. Nevertheless, female labour participation has marginally increased in this period.

Figure IV.2 Net labour participation, total, men and women, 2005-2011

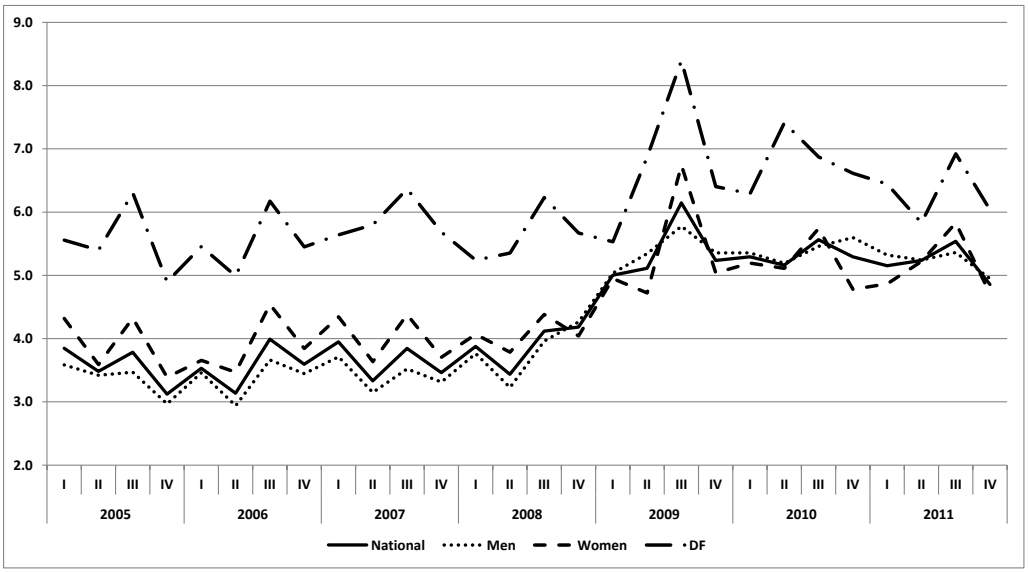


Source: own estimations using microdata from ENOE, 2011.

Figure IV.3 shows the quarterly unemployment rate[45] for men, women and national average as well as for the DF 2005 and 2011. Besides the observable seasonality due to business cycles; the 2008 financial crisis increased unemployment rates dramatically in 2009. Since around 2012, the gap between women and men narrows significantly.

45 The unemployment rate is defined as the ratio of the population looking for a job to the population economically active.

Figure IV.3 Unemployment rate, national, men, women and DF, 2005-2011



Source: own estimations using microdata from ENOE, 2011.

IV.2.3 Differences in gender and social classes

The Mexican society has some levels of inequality in wellbeing between gender and social classes. Table IV.1 shows differences in the equality of means on age, years of education, the divorced condition and working condition, between male and female population over 18 years old in 2012. This Table suggests a minimal difference between male and female in age and education, of 0.38 years on age greater on female over male and 0.30 years on education on male versus female. Regarding the percentage of divorced population, women have 3.8 percentage points more incidence than men, whereas men are participating in the labour market in 34.2 percentage points more than women.

Table IV.1 Test on the equality of means, men and women over 18 years old

Group	Age	Education	Divorced*	Working*
Male	37.34	9.36	3.13	84.22
	(0.0544)	(0.0184)	(0.0700)	(0.1400)
Female	37.73	9.07	6.94	49.99
	(0.0514)	(0.0175)	(0.0900)	(0.1900)
Difference	-0.3852	0.2961	-3.8093	34.2308
	(0.0748)	(0.0254)	(0.1200)	(0.2400)
$Pr(\|T\| > \|t\|)^A$	0.0000	0.0000	0.0000	0.0000

Source: own estimations using microdata from MCS-ENIGH, 2012.
[A] p-value associated with the equality of the mean across gender.
*Expressed in percentages.

Now, the same analysis focuses only on women, older than 18 years, who belongs to the deprived group or to the middle class as defined by the counting approach of chapter III. The middle class is 1.4 years older on average than the deprived. This group also has a higher education of almost four years. The deprived have greater rates of divorce, 1.3 percentage points more than women in the middle class, and work less, on 11.9 percentage points.

Table IV.2 Test on the equality of means, women deprived and middle class

Group	Age	Education	Divorced*	Working*
Deprived	36.83	6.57	7.74	42.51
	(0.0832)	(0.0231)	(0.1700)	(0.3100)
Middle class	38.23	10.54	6.44	54.41
	(0.0655)	(0.0212)	(0.1200)	(0.2400)
Difference	-1.4021	-3.9700	1.3016	-11.9015
	(0.1066)	(0.0329)	(0.2000)	(0.4000)
Pr(|T| > |t|)[A]	0.0000	0.0000	0.0000	0.0000

Source: own estimations using microdata from MCS-ENIGH, 2012.
[A] p-value associated with the equality of the mean across classes.
*Expressed in percentages.

Difference between women in the middle class and the affluent, over 18 years old, are also significant. Table IV.3 reports tests on these differences. On average, the affluent are 1.2 years older and have 3.2 years of education more than the middle class. The middle class have more incidence of divorce, of 2.6 percent points. Affluent women work more, 4.2 percent points than women in the middle class.

Table IV.3 Test on the equality of means, women affluent and middle class

Group	Age	Education	Divorced*	Working*
Affluent	39.27	13.33	4.21	58.02
	(0.1868)	(0.0578)	(0.2700)	(0.6800)
Middle class	38.10	10.12	6.79	53.83
	(0.0695)	(0.0219)	(0.1300)	(0.2600)
Difference	1.1682	3.2056	-2.5831	4.1966
	(0.1971)	(0.0621)	(0.3600)	(0.7400)
Pr(|T| > |t|)[A]	0.0000	0.0000	0.0000	0.0000

Source: own estimations using microdata from MCS-ENIGH, 2012.
[A] p-value associated with the equality of the mean across classes.
*Expressed in percentages.

IV.3 Literature review

This section presents a review of the literature on the social roles of household members, the intra-household decision-making process and the influence divorce has on it, as well as the attitudes of middle-class women to these aspects of social life. These discussions are presented in order to give theoretical foundations to further empirical analysis, as well as to interpret the results under the light of the theory.

IV.3.1 Social roles within households

Historically, the participation of women in the labour market has been limited in Latin America. However, this trend has changed in the last decades, particularly in urban areas where more women are joining the labour market (Beneria and Roldan, 1987; Chant, 1991). Smith and Ward (1985) explain that an increase in wage would lead to higher female labour market participation; however, empirical evidence shows this is not always true, and other economic, cultural and social factors could have been involved in this decision (Peters, 1986; Johnson and Skinner, 1986).

Cultural factors are relevant in the division of work within households. An example is found in the division of household responsibilities in Latin American households, where women are involved in household tasks with a higher level of responsibility since childhood, whereas men are involved in household activities with less level of responsibility (Valenzuela, 1999). While analysing social factors within working families, Hochschild and Machung (2012: p 15) find three types of ideologies on marital roles, which allows for household classifications: (1) traditional, in which male household members do extra-household work and women do housework; (2) transitional, where both household partners work outside home; and (3) egalitarian, where partners share housework in an egalitarian way.

INEGI reports that 87.8% of women do housework, whereas 12.2% of men do it (INEGI, 2000: p 24). However, extra-household work is done by 84.2% of men and nearly 50.0% of women (ENOE, 2011). Pedrero (2004) explains these gaps as differences on the household members' reproductive functions, and the social feminisation of household activities, making men less participative in housework. Then, based on intra-household working roles, Mexican households could be considered as traditional following Hochschild and Machung (2012).

IV.3.2 The intra-household decision-making process

Earlier household decision-making models consider a male household head in charge of the decision process in dictatorship fashion (Becker, 1976). This approach is also called the unitary decision-making model (Wheatley and Wu, 2014), and could be found in traditional households. This model, nevertheless, underestimates the individual preferences of other household members (Grossbard, 2010), and overestimates the decision roles rather than the decision processes and outcomes (Davis, 1976).

Alternative economic models analyse the household decision-making process as a collective household behaviour rather than unitary, where each household member is characterised by their particular preferences (Bourguignon and Chiappori, 1992). Another approach is based on cooperative games, where the bargaining between household members plays a crucial role in the decision-making process (Manser and Brown, 1980). The bargaining power is essential in the allocation of resources within households (Thomas, 1990; Luke and Munshi, 2006). The control of financial resources is also relevant for the decision-making process (Ashraf, 2009).

An important aspect of the bargaining process to consider is the differentiated role of male and female household members over nonwage income (Horney and McElroy, 1988). Income transfers between households do not challenge the decision-making control patterns within households by their own; other factors are required in order to influence the decision-making patterns like social intermediation (Holvoet, 2005).

Considering Mexican households as traditional, women have low bargaining power within the household decision-making process due to their lack of resources; therefore, women have reduced autonomy and freedom (CNDH, 2013). Thus, this is the main argument from the CNDH and the SCJN to introduce no-fault divorce.

IV.3.3 The role of divorce in the intra-household decision process

Divorce, as a legal symbol, is relevant for human freedoms (Marco, 2009), acting as a credible threat correlated to bargaining power within households (Sebastian, 2011). Attitudes towards divorce could be influenced by popular culture, for example through TV shows due to the exposition of modern lifestyles (e.g. soap operas portraying empowered female roles), of

emancipated female figures and the critic to traditional values (Chong and La Ferrara, 2009).

Arguably, no-fault divorce empowers women by providing an easier way to dissolve marriages without their partners' agreement, thus increasing women's bargaining power within households by giving a feasible menace of divorce, and to escape from problematic marriages (CNDH, 2013). However, no-fault divorce can also be used by men to leave the household easily, leaving women and children unprotected and vulnerable.

Deprived women in the USA feel difficult to escape from a difficult marriage (Miller and Sassler, 2011). Parkman (1992) found in the USA that when no-fault divorce is introduced with a narrow legal definition of property rights or limited alimony, female labour participation increased as well as their human capital accumulation.

Empirical studies in European countries show little or no effect of no-fault divorce on women's financial situation (Jacob, 1989). In the USA, no-fault divorce has no effect on female labour participation or in working time. However, it exerts a positive effect if one compares a married female with children to a married female without children (Genadek et al., 2007).

A possible explanation of the limited effect of the no-fault divorce introduction on women's labour participation in European countries or the USA could be the social benefits and institutional protection provided to the population in these countries, which are specially designed to protect vulnerable groups and to enhance their welfare. As found by Parkman (1992), when these benefits (law protection on property rights and alimony) are relaxed, then an effect of the divorce law introduction on women labour participation is observed. Therefore, a positive effect of the no-fault divorce introduction on women's labour participation is expected in developing countries with a lack of social coverage or high levels of informality, like Mexico.

In a comparative study of legal frameworks across Latin America, Marco-Navarro (2009) identifies divorce as a relevant legislation to increase human rights. Sebastian (2011) finds that divorce balances the bargaining power within household members. The Mexican Human Rights Commission considers that no-fault divorce legislation makes the process easy, affordable, and broadly accessible, as well as less traumatic for the people involved; therefore, this institution expects that its introduction would increase the opportunities of getting divorce particularly to women (CNDH, 2013).

In this way, The Mexican Supreme Court expects women to be empowered by the no-fault divorce due to an enhancement of their human rights, freedoms, and capabilities (SCJN, 2012). Therefore, an increase in women participation in the labour market would also be expected due to an enhanced women agency, or as a protection from vulnerability in case of being abandoned by their partners.

IV.3.4 The middle class

The use of divorce as a bargaining chip to improve women's conditions could be traced to Finland in the late 19th century, where women from the upper middle class used divorce as an instrument to promote changes in men's social behaviours, like drunkenness and domestic violence (Saarimaki, 2018). Cazenave (1983) highlights the strong relationship between race, gender and social class as determinants of social behaviours, and finds that in the USA, within the Afro-American middle-class community, men were more tolerant of domestic violence against women or children after being divorced, separated or remarried, compared to never-married men, widowers or married for the first time. This finding suggests that tensions inside the household could lead to violence even in educated, middle-class households.

As mentioned in subsection IV.2.3, women belonging to the middle class participate in the labour market on average more than women from deprived backgrounds (almost 12% more). However, as found by Hochschild and Machung (2012: p 19), women belonging to the middle class are in an advantageous position to participate more in the labour market due to progressive working policies in the workplace, generous benefits and high salaries allowing affordability for children daycare. These facilities are not always offered in jobs where women from lower social groups usually work.

Financial security provides women with more control over their own decisions. Asian women from the middle-class report that marriage is not observed as a financial aid, but as an imposition of constraints on personal freedoms (Williams and Guest, 2005). Therefore, Asian young middle-class women, usually economically independent, prefer to delay their marriages to prevent premature marital dissolution. In contrast, the average USA woman does not hesitate to marry for fear of divorce (Cherlin et al., 2008), but middle-class women have a greater commitment to marriage than women from deprived backgrounds (Miller and Sassier, 2011).

Despite the relevance of the middle class discussed in previous chapters, no studies were found analysing the effects of no-fault divorce on women from this social group, neither to its heterogeneous effects. This gap in the literature underpins the relevance of the findings documented in this chapter.

IV.4 Methodology

This section presents a discussion on the estimation methodologies to find out the impact of the no-fault divorce introduction. In the first part, a discussion on the different estimation approaches is presented, followed by two proposed methods in the existing literature; and finally with another method which is not being exploited before in the literature.

IV.4.1 The evaluation of social interventions

Different techniques and strategies to evaluate the impact of social interventions in a society are available in the public policy literature. One of the most critical issues faced in the application of those techniques consists of working with samples not designed *ad-hoc* for experimental purposes. Moreover, several challenges have arisen in an empirical analysis of this kind, for instance, the control for estimation biases like the identification problem (Manski, 1992: p 3), the consistency of individuals over time (Wooldridge, 2002: p 147), and on the selection-on-observables (Sianesi, 2014). The objective lies in evaluating the effect of the intervention programme regarding the estimate of the average causal effect of the intervening variable, usually binary, on the dependent one. The exposure to this causal variable could be defined as a treatment, and non-exposure as a control (Wooldridge, 2002: p 147).

The Difference-in-Difference (DID) method is the preferred estimation strategy for the analysis of natural quasi-experimental study cases, mainly when data was not designed *ad-hoc* as a controlled experiment following a randomised selection of control and treatment observations (Ashenfelter and Card, 1985; Angrist and Pischke, 2009: p 21). The DID method consists of comparing the variable of interest between treated and control individuals before and after the introduction of social interventions (Card and Krueger, 1994).

Ravallion et al. (2014: p 130) and Wooldridge (2002: p 148) explain how DID could be estimated as a difference on the means between groups rather than

individuals through non-panel repeated cross sections (RCS). However, misidentification of the groups between periods could lead to biased results (Greene, 2012: p 196), and standard errors in a double-difference regression are likely to be serially correlated across time (Bertrand et al., 2004). A solution to this issue could be to cluster standard errors and to work with them as fixed effects (Genadek et al., 2007).

The DID estimator is defined on a linear regression model expressed as:

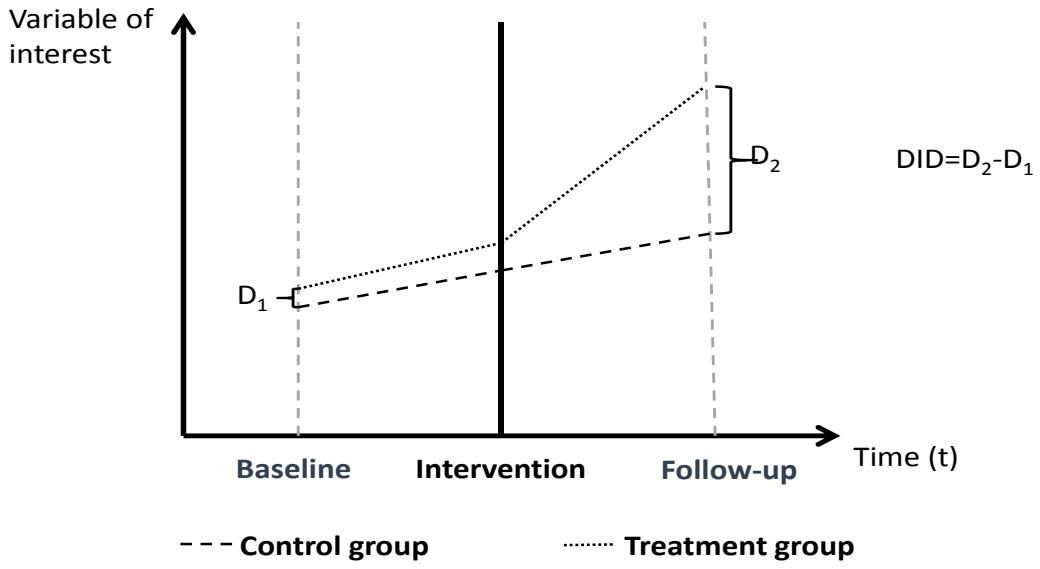
$$Y_{ist} = \alpha + \beta_1 time_{it} + \beta_2 treatment_{st} + \beta_3 time_{it} * treatment_{st} + \sum_q \beta_q \Phi_{istq} + \varepsilon_{ist} \quad \textbf{(IV.1)}$$

where Y_{ist} is the dependent variable for individual *i* in either treatment or control *s* condition at *t* time; *α* is the constant term; $time_{it}$ is a time dummy for individual *i* in the period *t* after the social intervention is introduced; $treatment_{st}$ is a dummy variable which identifies the treatment – τ – group from the control one – *c* –; $\sum_q \beta_q \Phi_{istq}$ stands for a matrix of *q* control variables; ε_{ist} is the error term; and $time_{it} * treatment_{st}$ represents the DID estimator, which is defined as:

$$\beta_3 = E[y_{i\tau 2} - y_{i\tau 1}] - E[y_{ic2} - y_{ic1}] \quad \textbf{(IV.2)}$$

Equation (IV.2) in terms of the coefficient (β_3) measures the effect of the social intervention as the difference of the estimated effect on the dependent variable before and after its introduction. A graphical representation of this estimation methodology is presented in Figure IV.4.

Figure IV.4 A generic graphical representation of a DID estimation



The gradual introduction of the no-fault divorce in Mexico has been followed by a random pattern independent of labour market conditions or individual

preferences[46]. These characteristics suggest a quasi-natural experimental study case (Genadek et al., 2007), where some sort of randomisation across individuals who are affected by the law's introduction (defined in this case as the treatment) is being guaranteed (Blundell and Costa-Dias, 2009).

True *ad-hoc* experimental neither panel data are available in Mexico. Then, a pooled cross-section analysis seems to be the most plausible alternative. Three estimation strategies are being suggested. The first one consists on the estimation of the average effect on the mean of the group of interest using an RCS method. The second one is an average treatment on the treated estimation based on the propensity score matching method. The latest one is a methodological contribution of this chapter, which builds a synthetic panel to estimate the average effect.

IV.4.2 First estimation strategy, the repeated cross-section analysis

The first DID estimation strategy comes from a simple linear regression specification as defined in equation (IV.1). When the dependent variable is a binary response, then the model has a *probit* specification defined as:

$$P(y_{it} = 1|x_{it}) = \alpha + \beta_1 time_{it} + \beta_2 treatment_{st} + \beta_3 time_{it} * treatment_{st} + \sum_q \beta_q \Phi_{istq} + \varepsilon_{ist}$$

(IV.3)

following the variable specification in equation (IV.1). The marginal effects, which report the probability of change of *x* on *y*, are computed as:

$$MgFx = \frac{\partial y}{\partial x} \quad \textbf{(IV.4)}$$

In the case of a continuous dependent variable, an ordinary least-squares linear regression is estimated following the same specification as equation (IV.1).

---

46 Some comments to this work argue that the no-fault divorce introduction in 2008 was not a strictly independent event from the society, as people elect their representatives based on their political views and expectative. However, the local assembly of representatives was elected in 2006, two years before the introduction of no-fault divorce, which was not included in the manifesto of any political party at the election.

IV.4.3 Second estimation strategy, the Propensity Score Matching

A second strategy for DID identification consists in the estimation of the Average Treatment effect on the Treated –ATT- in the non-panel cross-sectional context. The ATT effect is defined as:

$$ATT = E[y_{i2} - y_{i1} \mid \tau_i = 1] \quad \textbf{(IV.5)}$$

where ATT is the expected value of the difference in *y* before the treatment and after the treatment, under the condition of having received the treatment as $\tau_i = 1$. y_i for periods 1 and 2 are computed following equations (IV.1) or (IV.4) according to the nature of the dependent variable. The reason behind this estimation is because people could or could not be exposed to the treatment, meaning that it is not possible to calculate both conditions $\tau_i = 1$ on y_{i2} and y_{i1} for each individual. Therefore, a Propensity Score Matching (PSM) is estimated, which calculates a propensity score based on a group of covariates. This procedure matches individuals with similar scores but belonging to opposite treatment or control groups using the nearest neighbour approach, and then calculates the average treatment effect (ATT).

IV.4.4 Third estimation strategy, the synthetic panel

The third DID estimation strategy is a methodological contribution of this chapter, which assumes panel data conditions (same individuals appear in the baseline and follow up periods). In this way, fixed effects and unobserved heterogeneity are being controlled through maintaining the same sample of individuals before and after the social intervention (Greene, 2012: 388).

Considering the absence of panel data, the proposed estimation is based on synthetic panels, which are not widely used in DID estimations. Synthetic panels were introduced by Deaton (1985) and developed by Dang and Lanjouw (2013) as a method to analyse income poverty over time in countries with no availability of panel data. Synthetic panels consider the identification of population cohorts in cross-sectional periods based on socio-demographic characteristics. Once these population cohorts are identified, the information of interest from each cohort is collapsed by its average as follows:

$$I_k = \frac{\sum_{n=1}^{i} I_{ik}}{n_k} \quad \textbf{(IV.6)}$$

where I_k represents the average value of variable *I* within the cohort *k*; I_{ik} represents the value of the indicator *I* for individual *i* in the cohort *k*, and n_k represents the population in the cohort *k*.

The variance within cohorts is minimised through this technique and keeps cohorts homogenous over time. Thus, the original sample is re-grouped from individual observations into a synthetic panel of individual cohorts with values representing the mean within each cohort. In this case, the DID estimation follows the same functional form of equation (IV.1) or equation (IV.2), depending on the nature of the dependent variable. By substituting the right-hand side of equation (IV.6) in (IV.1), the estimation looks as follows:

$$\left(\frac{\sum_{n=1}^{i} Y_{ik}}{n_k}\right)_{kst} = \alpha + \beta_1 \left(\frac{\sum_{n=1}^{i} time_{ik}}{n_k}\right)_{kt} + \beta_2 \left(\frac{\sum_{n=1}^{i} treatment_{ik}}{n_k}\right)_{st} + \beta_3 \left(\frac{\sum_{n=1}^{i} time_{ik}}{n_k}\right)_{kt} * \left(\frac{\sum_{n=1}^{i} treatment_{ik}}{n_k}\right)_{st} + \sum_q \beta_q \Phi_{kstq} + \varepsilon_{kst} \quad \textbf{(IV.7)}$$

Thus, these cohorts keep the average information of similar individuals. Estimations between baseline and follow-up periods will match the same cohort before and after the social intervention (Angrist and Pischek, 2009: p 21). This matching of individuals (cohorts) with the same (or very close) information allows the control of the individuals' fixed effects over time.

Under certain circumstances, observable characteristics of the individuals could affect the conditional expectation on the treatment's effect, violating in this way the conditional independence assumption -the outcome for untreated observations should not affect their counterfactual condition (Greene, 2012: p 776).

These characteristics are known as selection-on-observables and could be used to establish a new control group (Sianesi, 2014) through a nonparametric Kernel estimator which uses local averaging to estimate a function at a particular point (Wooldridge, 2002: p 915). A PSM between treatment and control observations in the baseline period is calculated in terms of this Kernel adjusted estimator (Epanechnikov function), and the generated differences among treatment and control observations are used as weights in the DID estimation. This estimation will be referred in the empirical analysis as the Kernel synthetic panel estimation.

IV.5 Empirical analysis

This section reports the empirical analysis. The first three subsections introduces the data along with the challenges faced to identify the middle class and descriptive statistics. In the following subsections, the results are analysed using the three estimation approaches proposed in the previous section.

IV.5.1 Data source

Although Mexico has few panel data surveys available with sociodemographic information, those few available surveys are not designed to collect precise information about the labour market and have not enough statistical power for sample disaggregation at regional or state level. The collection frequency and seasonality of these surveys are also not suitable for the purposes of this analysis because they usually do not cover the periods before and after the no-fault divorce introduction.

This chapter uses the National Occupation and Employment Survey (ENOE). ENOE survey is a quarterly household survey conducted by the Mexican National Institute of Statistics and Geography (INEGI) since 2005, which collects information for all household members of 12 years or older, at State level representativeness. This survey is continuously collected along the year and helps to build the official indicators of labour market. Micro-level datasets of this survey are publicly available on quarterly basis.

IV.5.2 Identification of the middle class

A significant challenge faced in this analysis is the identification of the middle class in the ENOE sample. As explained in chapter III, in Mexico the ENIGH survey is the most comprehensive source of information on household income and expenditure, while the MCS-ENIGH survey is the best data source for the analysis of household characteristics.

The ENOE survey, on the other hand, is the main reference for labour information. However, this survey has not sufficient information either regarding household characteristics or income to identify the middle class as defined in chapters II and III. A limited set of variables in ENOE provide information on the household sociodemographic characteristics and its labour income. However, this income does not represent all other income sources[47], making unidimensional definitions of the middle class based on labour income biased.

Then, a matching identification strategy between the MCS-ENIGH 2012 and the ENOE is proposed in order to impute the middle-class status obtained by

47. Labour income is also sensitive to seasonality and labour market shocks like unemployment.

the counting approach in chapter III into ENOE households. Considering that both surveys are gathered by the same institution (INEGI), labour income, as well as a small set of sociodemographic characteristics, have the same definitions, guaranteeing comparability between the surveys.

The matching follows four steps:

1. An econometric linear estimation of a predicted per capita labour income in both surveys at household level, based on the set of comparable characteristics and the State where the household is located as a control.
2. Identification of middle-class thresholds (deprived and affluent) on the distribution of the predicted labour income in the MCS-ENIGH, using the middle-class sample identified by the counting approach in chapter III.
3. Application of these thresholds in the predicted per capita labour income distribution of ENOE 2005 and 2011[48], in the same manner as an absolute middle-class definition, using the predicted income as a wellbeing domain.
4. Assuming that all individual members of a middle-class household share this condition, then all individuals in an ENOE household identified as middle class are given this status.

The econometric model is based on a linear equation as follows:

$$y_h = \alpha + \sum_k \beta_k \Phi_h + \varepsilon_h \qquad \textbf{(IV.8)}$$

where y_h is the observed real income in monetary units for household *h*, α is the constant, ε_i the error term, and $\sum_k \beta_k \Phi_h$ represents a set of independent variables which are, for the household head gender, age, civil status and years of education; and the household demographic dependency rate, the household average number of working members, the urban/rural condition[49], and a set of dummy variables to control for the state where the household is located. Descriptive statistics on these variables are presented in Appendix C.1. The econometric estimations for each sample are in Appendix C.2.

Then, the predicted income is defined as:

$$\widehat{y_h} = \hat{\alpha} + \sum_k \widehat{\beta_k} \Phi_h \qquad \textbf{(III.9)}$$

---

48 In order to keep comparability between years, the predicted income was calculated in real terms in all samples, giving 2005 as the base year.

49 Cities are with more than 2,500 inhabitants.

where $\widehat{y_h}$ stands for the predicted household per capita labour income. Descriptive statistics for the predicted income in all samples, as well as the density plots, are given in Appendix C.3 and C.4 respectively.

Finally, the middle-class thresholds where identified on the predicted labour income ($\widehat{y_h}$) as \$0.0010 Mexican pesos for the deprived[50], and \$3,528.60 for the affluent. The estimated thresholds are then applied to the ENOE sample, obtaining an overall middle class (considering both 2005 and 2011 samples) of 1,850,520 individuals (85.5% of the sample). A plot of the predicted percentile distribution and thresholds is in Appendix C.5.

IV.5.3 Data sample

The analysis consists of all female respondents (to avoid sample selection bias), in working age (15 to 64 years old) belonging to the middle class. The four quarters within the years 2005 and 2011 (three years before and after the no-fault divorce introduction respectively) were merged and pooled to build a dataset covering the baseline and follow up periods.

In the case of the no-panel RCS estimations, the information from the pooled dataset was preserved at individual level. In the case of the synthetic panel, this sample was collapsed by the variable means of the population cohorts following equation (IV.6). The population cohorts' determinants are:

1. Quarter of the year, to control for seasonality.
2. Year of birth, to control for individuals' characteristics.
3. Federative State of residence, for treatment/control identification.

Both ENOE datasets (2005 and 2011) contain the same variables, and these were built identically. Considering age is an average within the cohort, there is a variance on age within these cohorts for those individuals whose birthday was after or before the survey's collection. Thus, age could be used in the econometric analysis as a control variable in the synthetic panel case, despite the year of birth was used as a determinant of the population cohort.

For all estimation strategies, the basic model is defined as equation (IV.1) or equation (IV.3) according to the dependent variable (dummies for women labour participation and labour informality, logarithm of worked hours). The

---

50 The value of this lower bound is close to zero because the calculation of the predicted labour income $\widehat{y_h}$ included several observations with value of zero, particularly in the ENOE sample.

explanatory variables are a set of sociodemographic characteristics including age, age squared, civil status (married or not), a dummy variable to identify if the individual is the household head, household size, the natural logarithm of the household real labour per capita income, levels of education achievement (in terms of the binary indicators such as no education, primary education, secondary education, higher secondary and professional, and post-graduate) household dependency rate of the young (0 to 14 years), and the elder (65 or more years), and their residence location (urban or not). The State where individuals are living were included as controls only in the first estimation strategy.

Table IV.4 shows the descriptive statistics of the variables included in the whole RCS estimations' sample for each baseline and follow up period. The female employment rate has increased from 41.4% to 45.8%. Informal employment grew more than 1 percent point. The natural logarithm of working hours in both periods stand at 4.0 and 4.6 worked hours per week for the baseline and follow up respectively. The average age has increased by one year, from 35.7 to 36.7. The percentage of middle-class women as household head has risen from 13.2 to 16.2, whereas the percentage of married population decreased by 0.4 percent points. The average household size is four individuals in both periods.

The logarithm of the real per capita income has decreased from 6.6 to 6.2, representing a reduction of $204.5 pesos in real terms. No education and primary education have reduced to 1.5 and 6.4 percent points, whereas education has improved in the levels of secondary, high school and professional education. The dependency rate on young population has decreased by 0.1 units, whereas the dependency rate on elderly population remains almost the same. Urban population has decreased by 2 percent points. Finally, the proportion of population living in Mexico City (DF) is 2.7% on average.

Table IV.4. Descriptive statistics, RCS, 2005-2011

Variable	Observations	Mean	Estandard deviation	Minimum	Maximum
Baseline					
Employed*	310,822	41.4	49.3	0.0	100.0
Informal employment*	128,807	55.5	49.7	0.0	100.0
Working hours (logarithm)	310,822	1.4	1.8	0.0	5.0
Age	310,822	35.7	13.2	15.0	64.0
Age square	310,822	1,446.0	1,002.0	225.0	4,096.0
HH head*	310,822	13.2	33.9	0.0	100.0
Married*	310,822	59.7	49.0	0.0	100.0
HH size	310,822	4.1	1.7	1.0	21.0
HH per capita income (logarithm)	310,822	6.6	2.2	0.0	12.2
No education*	310,822	5.2	22.1	0.0	100.0
Basic education*	310,822	31.4	46.4	0.0	100.0
Secondary education*	310,822	27.5	44.7	0.0	100.0
High school and professional*	310,822	22.0	41.4	0.0	100.0
Posgraduated*	310,822	13.9	34.6	0.0	100.0
HH dep rate, young	310,822	0.40	0.52	0.00	7.00
HH dep rate, old	310,822	0.07	0.24	0.00	5.00
Urban*	310,822	86.7	34.0	0.0	100.0
Mexico City*	310,822	2.9	16.8	0.0	100.0
Follow up					
Employed*	330,631	45.8	49.8	0.0	100.0
Informal employment*	151,544	56.6	49.6	0.0	100.0
Working hours (logarithm)	330,631	1.5	1.8	0.0	5.1
Age	330,631	36.7	13.5	15.0	64.0
Age square	330,631	1,528.0	1,034.1	225.0	4,096.0
HH head*	330,631	16.2	36.9	0.0	100.0
Married*	330,631	59.3	49.1	0.0	100.0
HH size	330,631	3.9	1.6	1.0	16.0
HH per capita income (logarithm)	330,631	6.2	2.5	0.0	12.0
No education*	330,631	3.7	18.9	0.0	100.0
Basic education*	330,631	25.0	43.3	0.0	100.0
Secondary education*	330,631	29.2	45.5	0.0	100.0
High school and professional*	330,631	24.1	42.8	0.0	100.0
Posgraduated*	330,631	18.0	38.4	0.0	100.0
HH dep rate, young	330,631	0.34	0.48	0.00	6.00
HH dep rate, old	330,631	0.07	0.24	0.00	4.00
Urban*	330,631	84.7	36.0	0.0	100.0
Mexico City*	330,631	2.6	16.0	0.0	100.0

Source: own estimations using microdata from ENOE, 2005-2011.
*Numbers in percentages.

Table IV.5 shows descriptive statistics of the synthetic panel dataset for the baseline and follow-up periods. The female employment rate is 43.6% and 49.3% respectively, while informal employment stands at 54.5% and 54.9%. The natural logarithm of working hours is on average of 4 and 5 worked hours per week in baseline and follow-up periods respectively. Average age has increased from 34.5 to 40.5. Female household heads has increased from 13.2% to 19.8%, whereas the percentage of married population moved from 59.6% to 65.6% respectively. The average household size is between 4.7 and 4.4 individuals in both periods. The natural logarithm of the real per capita income has decreased from 6.5 to 6.1 representing a real decrease of $224.4 pesos. Levels of education show the same pattern as in the case with RCS sample. The dependency rate on young population has decreased during this

period from 0.52 to 0.44, whereas adult dependency rate increased from 0.07 to 0.09. Urban population has decreased by 0.6 percent point. Finally, the proportion of population living in Mexico City (DF) remains the same across the period.

Table IV.5 Descriptive statistics, synthetic panels, 2005-2011

Variable	Observations	Mean	Estandard deviation	Minimum	Maximum
Baseline					
Employed*	4,126	43.6	13.6	0.0	74.0
Informal employment*	4,124	54.5	16.8	5.9	100.0
Working hours (logarithm)	4,126	1.5	0.5	0.0	2.6
Age	4,126	34.5	11.9	15.0	58.0
Age square	4,126	1,328.7	843.3	225.0	3,364.0
HH head*	4,126	13.2	9.9	0.0	54.5
Married*	4,126	59.6	23.5	0.0	93.1
HH size	4,126	4.7	0.5	2.9	6.4
HH per capita income (logarithm)	4,126	6.5	0.4	4.1	7.5
No education*	4,126	4.7	5.6	0.0	48.5
Basic education*	4,126	29.5	14.7	1.0	79.2
Secondary education*	4,126	27.1	13.7	0.0	94.6
High school and professional*	4,126	22.5	11.5	0.0	71.8
Posgraduated*	4,126	16.3	9.2	0.0	48.6
HH dep rate, young	4,126	0.52	0.26	0.03	1.30
HH dep rate, old	4,126	0.07	0.03	0.00	0.29
Urban*	4,126	85.7	5.6	63.8	100.0
Mexico City*	4,126	3.7	18.8	0.0	100.0
Follow up					
Employed*	4,126	49.3	11.0	8.8	77.0
Informal employment*	4,126	54.9	14.8	12.8	100.0
Working hours (logarithm)	4,126	1.7	0.4	0.3	2.8
Age	4,126	40.5	11.9	21.0	64.0
Age square	4,126	1,778.2	984.2	441.0	4,096.0
HH head*	4,126	19.8	10.8	0.0	60.0
Married*	4,126	65.6	12.1	20.9	87.8
HH size	4,126	4.4	0.5	2.8	6.0
HH per capita income (logarithm)	4,126	6.1	0.6	3.0	7.6
No education*	4,126	4.7	5.4	0.0	42.4
Basic education*	4,126	28.1	14.3	0.0	76.0
Secondary education*	4,126	25.4	8.5	0.0	54.2
High school and professional*	4,126	21.0	7.7	0.0	51.6
Posgraduated*	4,126	20.8	9.2	0.0	54.7
HH dep rate, young	4,126	0.44	0.24	0.04	1.12
HH dep rate, old	4,126	0.09	0.06	0.00	0.55
Urban*	4,126	85.1	5.8	54.3	100.0
Mexico City*	4,126	3.7	18.8	0.0	100.0

Source: own estimations using microdata from ENOE, 2005-2011.
*Numbers in percentages.

IV.5.4 Estimation of the no-fault divorce effects on labour participation

Table IV.6 shows the estimates for the covariates' coefficients for all the three DID estimations on women labour participation. Column (1) and (2) corresponds to the first method (*probit*, where dependent variable equal to 1

if woman working, and 0 otherwise) with bootstrapped and robust standard errors respectively; column (3) for the second estimation method; column (4) and (5) for the synthetic panel direct estimation and the Kernel adjusted estimation respectively. Note that since some individuals were followed during different quarters across the sampling period, standard errors were estimated using bootstrapping with 100 replications (in order to account for violations of the assumption of independent distributions).

Table IV.6 Female labour market participation

Variables	(1)	(2)	(3)	(4)	(5)
Age	0.0634***	0.0634***	0.0484***	0.045***	0.8217***
	(0.0002)	(0.0002)	(0.0053)	(0.001)	(0.1847)
Age square	-0.0007***	-0.0007***	-0.0003***	-0.001***	-0.0071***
	(0.0000)	(0.0000)	(0.0000)	(0.000)	(0.0022)
Married	-0.2264***	-0.2264***	-0.1485***	-0.08***	-10.550***
	(0.0013)	(0.0015)	(0.0384)	(0.013)	(2.0779)
HH head	0.1600***	0.1600***	-0.3698***	0.258***	-10.753***
	(0.0020)	(0.0020)	(0.0301)	(0.019)	(2.9274)
HH size	-0.0036***	-0.0036***	0.0031	-0.007***	-0.5371*
	(0.0004)	(0.0003)	(0.0077)	(0.0030)	(0.3241)
Percapita income	0.0549***	0.0549***	-0.0551***	0.042***	0.0615
(logarithm)	(0.0002)	(0.0002)	(0.0039)	(0.0020)	(0.2519)
No education	-0.1352***	-0.1352***	-1.0414***	-0.119***	-15.351***
	(0.0035)	(0.0032)	(0.0818)	(0.0200)	(5.0455)
Primary education	-0.1236***	-0.1236***	-0.6406***	-0.2730***	-11.321***
	(0.0015)	(0.0018)	(0.0373)	(0.0140)	(2.1575)
Secondary	-0.0967***	-0.0967***	-0.0845***	-0.328***	-1.3506
education	(0.0018)	(0.0017)	(0.0325)	(0.0120)	(1.1678)
High school and	-0.0689***	-0.0689***	0.1775***	-0.196***	-0.8679
professional	(0.0019)	(0.0017)	(0.0312)	(0.0120)	(1.2287)
HH dependency	0.0405***	0.0405***	-0.2354***	0.0350***	-0.1313
rate, young	(0.0013)	(0.0012)	(0.0286)	(0.0060)	(0.8454)
HH dependency	0.1590***	0.1590***	-0.0187	0.2560***	8.6595***
rate, old	(0.0028)	(0.0027)	(0.0420)	(0.022)	(3.0940)
Urban	0.0163***	0.0163***	1.1500***	0.0760***	40.959***
	(0.0014)	(0.0016)	(0.0553)	(0.0150)	(3.7154)
Time	0.0442***	0.0442***	-	-	-
	(0.0010)	(0.0011)			
Constant	-	-	-5.7824***	-	-45.625***
			(0.1136)		(4.3066)
Controls by State	Yes	Yes	No	No	No
Controls by industry	Yes	Yes	No	No	No
N	635,269	635,269	635,269	8,252	8,252
Pseudo R^2	0.1733	0.1733	0.0280	0.7300	-
Bootstrapping reps	100	-	-	100	100

Source: own estimations using microdata from ENOE, 2005-2011.
Note: Standard errors in brackets; levels of significance at 90%, 95% and 99% as *, ** and *** respectively.

Regarding the covariates, these results seem consistent and robust across the estimation methods except in case of Column (3), and this is possibly due to the propensity score matching weights (PSM).

Given the focus on female labour market participation, Table IV.7 shows the effects of the no-fault divorce introduction on women labour participation as an interaction effect. This is being captured by β_3 as in equation (IV.3), or in equation (IV.5) or in equation (IV.7). Note in case of the *probit* model, the marginal effect associated with the estimated coefficient in equation (IV.3) is being reported. The first two columns refer to the marginal effects associated with the estimated $\hat{\beta}_3$ as in equation (IV.3) (with bootstrapped and robust

standard errors respectively). Column 3 of Table IV.7 reports the ATT difference estimate after PSM, whereas Columns 4 and 5 report the synthetic panel and Kernel adjusted estimator (equation (IV.7)).

Table IV.7 Impact of no-fault divorce on women labour market participation estimations

Interaction effect BS	Interaction effect RSE	ATT Difference	Synthetic panel	Synthetic panel K
0.0350***	0.0350***	0.0429***	0.0250***	0.0540
(0.0070)	(0.0067)	(0.0054)	(0.0070)	(0.0350)

Source: own estimations using microdata from ENOE, 2005 and 2011.
Note: Standard errors in brackets; levels of significance at 90%, 95% and 99% as *, ** and *** respectively.

These estimators share the same sign and significance level, suggesting a positive effect of no-fault divorce on the probability of women labour market participation between 2.5 and 3.5 percent points. In other words, this implies that the probability of participation in the labour market would increase on average by 2.5-3.5 percent points for a woman getting exposed to no-fault divorce compared to the one who cannot exercise this option. The average effect of exposition to no-fault (ATT) divorce is nearly 4.3 percent points on the probability of women labour market participation. The Kernel adjusted estimator has a higher coefficient compared to the other estimations, and it is not statistically significant at any level.

IV.5.5 Estimation of the no-fault divorce effects on informal labour market participation

Table IV.8 is similar as Table IV.6, presenting the estimates for the covariates for the three DID estimation strategies on women participation in the informal labour market. Column (1) and (2) refers to the first strategy (*probit*, 1 woman participating in the informal labour market, and 0 otherwise), with bootstrapped standard errors without and with industry dummies[51] respectively; column (3) and (4) are the same estimations with robust standard errors; column (5) and (6) the ATT-PSM estimation strategy without and with industry dummies respectively; column (7) the synthetic panel; and column (8) the Kernel adjusted estimator.

51 At International System of Industrial Classification, SIC-2 digit level of aggregation.

Table IV.8 Women's informal labour market participation

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Age	-0.0284***	-0.0206***	-0.0284***	-0.0206***	0.0561***	0.0550***	-0.0200***	0.8217***
	(0.0005)	(0.0003)	(0.0004)	(0.0004)	(0.0095)	(0.0096)	(0.0020)	(0.1847)
Age square	0.0003***	0.0002***	0.0003***	0.0002***	-0.0003***	-0.0003***	0.0000***	-0.0071***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0001)	(0.0001)	(0.000)	(0.0022)
Married	0.0771***	0.0668***	0.0771***	0.0668***	-0.0923*	-0.0893*	-0.133***	-10.550***
	(0.0021)	(0.0018)	(0.0021)	(0.0018)	(0.0498)	(0.0498)	(0.028)	(2.0779)
HH head	-0.0050**	-0.0044*	-0.0050*	-0.0044*	-0.2868***	-0.2551***	0.197***	-10.753***
	(0.0025)	(0.0023)	(0.0027)	(0.0023)	(0.0409)	(0.0410)	(0.0330)	(2.9274)
HH size	-0.0062***	-0.0042***	-0.0062***	-0.0042***	-0.0246**	-0.0215*	0.0580***	-0.5371*
	(0.0005)	(0.0004)	(0.0005)	(0.0004)	(0.0113)	(0.0113)	(0.0050)	(0.3241)
Percapita income (logarithm)	-0.0184***	-0.0123***	-0.0184***	-0.0123***	-0.0938***	-0.0940***	-0.0040	0.0615
	(0.0005)	(0.0004)	(0.0005)	(0.0004)	(0.0068)	(0.0069)	(0.0030)	(0.2519)
No education	0.5959***	0.2733***	0.5959***	0.2733***	-0.9445***	-0.9411***	0.9830***	-15.351***
	(0.0058)	(0.0054)	(0.0059)	(0.0057)	(0.1214)	(0.1262)	(0.040)	(5.0455)
Primary education	0.4697***	0.1842***	0.4697***	0.1842***	-0.4733***	-0.4798***	0.5540***	-11.321***
	(0.0023)	(0.0030)	(0.0022)	(0.0026)	(0.0513)	(0.0603)	(0.0240)	(2.1575)
Secondary education	0.3003***	0.0837***	0.3003***	0.0837***	-0.0504	-0.0365	0.4720***	-1.3506
	(0.0025)	(0.0026)	(0.0022)	(0.0024)	(0.0459)	(0.0527)	(0.0240)	(1.1678)
High school and professional	0.1471***	0.0101***	0.1471***	0.0101***	0.2182***	0.2153***	0.3220***	-0.8679
	(0.0026)	(0.0023)	(0.0024)	(0.0023)	(0.0433)	(0.0463)	(0.0250)	(1.2287)
HH dependency rate, young	0.0316***	0.0280***	0.0316***	0.0280***	-0.1686***	-0.1568***	-0.0320***	-0.1313
	(0.0017)	(0.0016)	(0.0017)	(0.0015)	(0.0373)	(0.0373)	(0.0090)	(0.8454)
HH dependency rate, old	-0.0223***	-0.0134***	-0.0223***	-0.0134***	-0.0545	-0.0488	0.3140***	8.6595***
	(0.0029)	(0.0027)	(0.0031)	(0.0028)	(0.0545)	(0.0544)	(0.0380)	(3.0940)
Urban	-0.1189***	-0.1294***	-0.1189***	-0.1294***	0.9197***	0.7888***	-0.3560***	40.959***
	(0.0027)	(0.0024)	(0.0028)	(0.0025)	(0.0797)	(0.0800)	(0.0270)	(3.7155)
Time	0.0390***	0.0192***	0.0390***	0.0192***	-	-	-	-
	(0.0016)	(0.0015)	(0.0017)	(0.0015)				
Constant	-	-	-	-	-5.5247***	-7.3606***	-	-45.625***
					(0.2048)	(0.4529)		(4.3067)
Controls by State	Yes	Yes	Yes	Yes	No	No	No	No
Controls by industry	No	Yes	No	Yes	No	No	No	No
N	274,398	274,398	274,398	274,398	274,398	274,398	8,250	8,250
Pseudo R^2	0.1847	0.3384	0.1847	0.3384	0.0236	0.0320	0.4800	-
Bootstrapping reps	100	100	-	-	-	-	100	100

Source: own estimations using microdata from ENOE, 2005 and 2011.
Note: Standard errors in brackets; levels of significance at 90%, 95% and 99% as *, ** and *** respectively.

Table IV.9 presents the effects of the no-fault divorce introduction on women participation in the informal labour market. This is being captured by β_3 as in equation (IV.3), or in equation (IV.5) or in equation (IV.7).

Table IV.9 Impact of no-fault divorce on women's informal labour market participation

Interaction effect BS	Interaction effect BS^C	Interaction effect RSE	Interaction effect RSE^C	ATT Difference	ATT Difference^C	Synthetic panel	Synthetic panel K
-0.0325***	-0.0201***	-0.0325***	-0.0201**	-0.0321***	-0.0317***	-0.0250**	-0.0430
(0.0098)	(0.0078)	(0.0099)	(0.0088)	(0.0078)	(0.0078)	(0.0110)	(0.0590)

Source: own estimations using microdata from ENOE, 2005 and 2011.
Note: Standard errors in brackets; levels of significance at 90%, 95% and 99% as *, ** and *** respectively. ^{C} Estimations controlled by industrial dummies.

All these estimations are negative and statistically significant at 5%, except for the Kernel estimation. The size of the estimated effects remains almost the same across estimation methods. These results suggest that the no-fault divorce introduction has a negative effect between 2.0 and 3.3 percent points on the probability of women informal labour market participation. A negative effect is also found in the average effect on women exposed to no-fault divorce at 3.2 percent points. Thus, when divorce becomes easy women seek for social security protection.

IV.5.6 Estimation of the no-fault divorce effects on working hours

Table IV.10 shows the estimated coefficients associated with the natural logarithm of the number of hours worked, which is used as the dependent variable. Column (1) and (2) refer to the OLS estimation of the first estimation method without and with industry dummies with bootstrapped standard errors; (3) and (4) are the same but with robust standard errors. Column (5) and (6) reports the results from the second estimation method (PSM) without and with industry dummies respectively; whereas column (7) contains the synthetic panel; and column (8) the Kernel adjusted estimator for the synthetic panel.

In columns (1)-(6), in order to control for the presence of outliers, the number of working hours is being restricted to lie in between 0 and 100 hours on a weekly basis. In columns (7)-(8), this restriction is being removed given that there are very few observations where the number of working hours is higher than 100 hours. **Table IV.10** Number of hours worked

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Age	0.0166***	0.0182***	0.0166***	0.0182***	0.0541***	0.0530***	0.153***	0.8217***
	(0.0007)	(0.0007)	(0.0007)	(0.0007)	(0.0097)	(0.0098)	(0.004)	(0.1847)
Age square	-0.0002***	-0.0002***	-0.0002***	-0.0002***	-0.0003***	-0.0003***	-0.002***	-0.0071***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0001)	(0.0001)	(0.000)	(0.0022)
Married	0.0297***	0.0303***	0.0297***	0.0303***	-0.0823	-0.0793	-0.289***	-10.550***
	(0.0038)	(0.0036)	(0.0036)	(0.0035)	(0.0510)	(0.0510)	(0.048)	(2.0779)
HH head	-0.1591***	-0.1572***	-0.1591***	-0.1572***	-0.2798***	-0.2473***	0.852***	-10.753***
	(0.0027)	(0.0028)	(0.0028)	(0.0028)	(0.0419)	(0.0420)	(0.067)	(2.9274)
HH size	0.0126***	0.0130***	0.0126***	0.0130***	-0.0267**	-0.0237**	-0.028***	-0.5371*
	(0.0008)	(0.0007)	(0.0007)	(0.0007)	(0.0115)	(0.0116)	(0.008)	(0.3241)
Percapita income (logarithm)	0.0227***	0.0235***	0.0227***	0.0235***	-0.0914***	-0.0914***	0.118***	0.0615
	(0.0007)	(0.0006)	(0.0006)	(0.0006)	(0.0071)	(0.0071)	(0.006)	(0.2519)
No education	-0.0918***	-0.0686***	-0.0918***	-0.0686***	-0.9692***	-0.9572***	-0.313***	-15.351***
	(0.0071)	(0.0079)	(0.0077)	(0.0080)	(0.1240)	(0.1287)	(0.075)	(5.0455)
Primary education	-0.0315***	-0.0210***	-0.0315***	-0.0210***	-0.4879***	-0.4865***	-0.862***	-11.321***
	(0.0032)	(0.0044)	(0.0034)	(0.0041)	(0.0526)	(0.0616)	(0.054)	(2.1575)
Secondary education	0.0215***	0.0129***	0.0215***	0.0129***	-0.0620	-0.0398	-1.107***	-1.3506
	(0.0031)	(0.0037)	(0.0031)	(0.0035)	(0.0472)	(0.0539)	(0.041)	(1.1678)
High school and professional	0.0605***	0.0305***	0.0605***	0.0305***	0.2082***	0.2140***	-0.659***	-0.8679
	(0.0029)	(0.0031)	(0.0029)	(0.0030)	(0.0446)	(0.0476)	(0.046)	(1.2287)
HH dependency rate, young	-0.0715***	-0.0689***	-0.0715***	-0.0689***	-0.1590***	-0.1472***	0.0910***	-0.1313
	(0.0026)	(0.0029)	(0.0026)	(0.0025)	(0.0380)	(0.0380)	(0.020)	(0.8454)
HH dependency rate, old	0.0451***	0.0422***	0.0451***	0.0422***	-0.0644	-0.0580	0.823***	8.6595***
	(0.0037)	(0.0034)	(0.0039)	(0.0039)	(0.0561)	(0.0560)	(0.072)	(3.0940)
Urban	0.1064***	0.1112***	0.1064***	0.1112***	0.9240***	0.7943***	0.3530***	40.959***
	(0.0042)	(0.0040)	(0.0044)	(0.0045)	(0.0814)	(0.0817)	(0.045)	(3.7155)
Time	-0.0175***	-0.0158***	-0.0175***	-0.0158***	-	-	-	-
	(0.0022)	(0.0022)	(0.0024)	(0.0023)				
Constant	3.1176***	3.1072***	3.1176***	3.1072***	-5.4955***	-7.3132***	-	-45.625***
	(0.0152)	(0.0183)	(0.0159)	(0.0174)	(0.2088)	(0.4546)		(4.3067)
Controls by State	Yes	Yes	Yes	Yes	No	No	No	No
Controls by industry	No	Yes	No	Yes	No	No	No	No
N	262,447	262,447	262,447	262,447	262,460	262,460	8,250	8,250
Adjusted R^2/$R^{2\,A}$	0.0541	0.0909	0.0542	0.0912	0.0235	0.0321	0.700	-
Bootstrapping reps	100	100	-	-	-	-	100	100

Source: own estimations using microdata from ENOE, 2005 and 2011.
Note: Standard errors in brackets; levels of significance at 90%, 95% and 99% as *, ** and *** respectively. $^{A}R^2$ reported in the case of the robust SE estimation.

Results seem similar among the different estimation strategies regarding significance levels, size of the estimated coefficients and signs. However, similar behaviour in terms of the size, sign and significance of the estimated coefficient is found as in the previous two cases, where either women labour force participation or women labour force participation in the informal sector is being modelled, for the PSM estimation.

Finally, Table IV.11 reports the marginal effects of the no-fault divorce introduction on the natural logarithm of women working hours for each estimation strategy as defined before. These results in terms of the size and significance remain similar except in the case of the synthetic panel. These results show a positive effect of no-fault divorce introduction on women working hours of more than 4.0% (1 more hour per week on average).

Table IV.11 Women's labour working hours, marginal effect estimations

Interaction effect BS	Interaction effect BS^{C}	Interaction effect RSE	Interaction effect RSE^{C}	ATT Difference	ATT Difference^{C}	Synthetic panel	Synthetic panel K
0.0457***	0.0414***	0.0457***	0.0414***	0.0360***	0.0361***	0.0770***	0.1970
(0.0137)	(0.0124)	(0.0132)	(0.0129)	(0.0091)	(0.0091)	(0.0260)	(0.1240)

Source: own estimations using microdata from ENOE, 2005 and 2011.
Note: Standard errors in brackets; levels of significance at 90%, 95% and 99% as *, ** and *** respectively. ^{C} Estimations controlled by industrial dummies.

A common critique of DID estimations is the failure to identify if the effect on the variable of interest is a result of the social intervention or if the effect is an independent shift in trends (Angrist and Pischke, 2009: p 52; Bertrand et al., 2004). Corrections for selection bias or the estimation of efficient standard errors are not sufficient measures to address this issue.

A practical alternative to test this problem is through the assumption of equal trends. This assumption suggests that under normal conditions, without the social intervention, trends of the variable of interest for treatment and control groups should remain equal. A way to operationalise a test for this assumption is by the estimation of placebos, consisting of the DID estimation on "fake" treatment groups which are known for not being affected by the social intervention (Gertler et al., 2011: p 101).

Placebo estimations were conducted for this analysis for the periods 2005-2007 (before the no-fault divorce introduction) and 20015-2017 (after the universalisation of no-fault divorce). Considering the consistency observed between RCS estimations and the synthetic panel, placebo effects were calculated exclusively for RCS samples due to computational simplicity. Placebo estimations show no effects, except for participation in labour market, suggesting equality of trends in both periods for informal labour and working hours. These placebos are reported in Appendix C.6.

IV.6 Conclusion

The objective of this chapter is to assess the impact of the introduction of no-fault divorce in Mexico on the labour market participation of middle-class women. This is being captured in terms of three variables: women's labour market participation, labour market participation in the informal sector, and the number of hours worked. The empirical estimation consists of assessing these effects using three DID estimations strategies: two non-panel RCS widely used in the literature, and the one proposed as an original contribution based on synthetic panels.

The results show that on an average the probability of participation in the labour market increases by 3.5 percent points whereas that in the case of informal labour market decreases by 3.1 percent points. An increase in working hours of 1 hour on average is being observed.

Regarding the estimation strategies, DID estimations based on the synthetic panel yield similar coefficients along with the almost similar significance level in comparison to those obtained by no-panel RCS. This could occur since both estimators, RCS and synthetic panel have used the average values of the same dataset. However, synthetic panels have a natural control on fixed effects and less variance within the variables, making them more efficient.

On the other hand, the Kernel adjusted DID estimator overestimates the social intervention effect, particularly in the case of the continuous dependent variable, the natural logarithm of working hours. A possible explanation could be that synthetic panels are measured in terms of the cohorts' means; therefore, the influence of selection-on-observables on the conditional expectation of the treatment effect is being minimised. In other words, if the conditional independence assumption of the control group is not violated; then, the Kernel adjusted estimators overestimate the DID coefficient. A disadvantage of using synthetic panels for DID estimation is the loss of degrees of freedom compared to RCS.

Future research could explore other RCS techniques and compare the effectiveness and efficiency of DID estimations based on synthetic panel data, like PSM with different settings of the nearest neighbour approach, radius matching as an alternative to the Kernel matching for instance, as well as exploring different ways to generate the population cohorts.

One of the main differences between this work and similar studies in developed countries like Jacob (1989) and Genadek et al. (2007), is data availability. The works by Jacob and Genadek et al. have used census data

covering a timeframe of decades and thereby suggesting long-term effects, whereas the analysis in this chapter relies on a three three-year period before and after the legal innovation. Further research could re-estimate these effects using the 2000 and the forthcoming 2020 censuses. Another outcome from this analysis is the possibility of observing heterogeneous effects as interaction of the treatment on the independent variables.

Chapter V
The real cost of financial exclusion among the middle class

V.1 Introduction

This chapter analyses the real interest rate differential paid for informal credit between financially excluded middle-class individuals and those not excluded. The empirical analysis focuses on urban middle-class individuals, but also provides results for those classified as deprived and affluent for comparative purposes.

Financial exclusion could prevent people from functioning correctly in a society (European Commission, EC, 2008: p 9). Moreover, credit is one of the most important resources to afford capital assets and properties (OECD, 2010); smoothing consumption over time and protecting households against adverse shocks (Zeller et al., 1997: p 12). The middle class provides an essential pool of consumers to the economy who are willing to pay extra for better quality products (Banerjee and Duflo, 2008), and consumes more than other classes (Wagley. 2013). Then, a creditworthy middle class is important for the economy.

This chapter contributes to the existing literature in several ways: first, no previous analysis has focused on the real interest rate differential between social classes. Second, there are no previous studies using information on both amounts borrowed and repaid from informal loans, which enable calculation of the real interest rate and, finally, this chapter considers the role of trust between individuals as a control for information screening, which is not common in the existing literature.

The main research question in this chapter is the identification of the real interest rate differential for informal loans between middle-class individuals with and without formal financial services, which are defined as financially included in the former case and excluded in the later. This proxy for financial exclusion follows constraints on data availability. However, it could be controversial as it does not differentiate between financial exclusion and non-users of formal financial services. A follow-up research question naturally emerges: identification of these differences within the other social classes.

Previous studies have found that social networks (the set of social relations an individual has) could help tame market information asymmetries. Bloom et al. (2008) argue that trust on social networks could overcome information asymmetries. Lin et al. (2009; 2013) found that the borrower's social network

could work as a screening mechanism to improve her probability to get a loan at a lower interest rate. However, crime and violence could damage social trust by destroying social cohesion, making trade more difficult (Moser and Holland, 1997; Colletta and Cullen, 2000). Mexico experienced in recent years an increase in crime, mostly related to drug cartels (between 2006 and 2012, the number of drug-related homicide grew 148.8%). Therefore, this analysis controls for crime conditions.

The methodology is based on a quasi-experimental framework where individuals financially excluded are defined as treatments, and non-excluded as controls. Thus, treatment effects would be estimated on the condition of exclusion. Differences are computed based on a propensity score matching, which pairs individuals with similar probability of being financially excluded, controlled by the loan amount. Once individuals are matched, the differences in real interest rates between treatment and control groups are computed in the form of the average treatment effect (between excluded and non-excluded), the average treatment on the treated (within the excluded), and the average treatment on the not-treated (within the non-excluded).

Results show that financially excluded middle-class individuals pay 7.3 percent points more for informal credit compared to non-excluded individuals, controlling for sociodemographic characteristics, loan amount, crime in the municipality where they live, and trust on the borrower's social networks and institutions. This difference could be interpreted as the cost of financial exclusion. In the case of the deprived and the affluent, the difference in interest rates between excluded and non-excluded individuals is in general statistically insignificant.

This chapter is divided in six sections. The following section gives a general view of the situation of financial exclusion and crime in Mexico. Section three presents the literature review on financial exclusion, credits and information asymmetries, social networks and social capital, and on how social capital could control information asymmetries. Section four discusses the methodology. Section five describes the data, the middle-class identification, and the results. Finally, section six concludes.

V.2 Context

This section presents an overview of the Mexican context in the dimensions of financial exclusion, crime and violence. This information would help to frame the results of the empirical analysis, as well as methodological decisions on the empirical analysis, within the socio-economic context in 2011.

V.2.1 Financial exclusion in Mexico

Financial exclusion is a relevant topic of economic development in Mexico. Access to financial services like savings accounts, credit and insurance, allows people to smooth their basic consumption over time, particularly when facing adverse shocks like accidents, unemployment or chronical illness (Zeller et al., 1997: p 12). Moreover, credit access allows people to consume or acquire capital goods, assets and properties which are unaffordable otherwise (Lyons, 2003; OECD, 2013).

Therefore, financial exclusion is considered an undesirable condition and a central concern for social development policy. The Mexican Ministry of Finance (SHCP) argues that, considering the low participation of workers in the formal economy[52], a significant share of people is economically vulnerable without coverage of social security and pensions. Then, financial inclusion is critical for social wellbeing (SHCP, 2013: p 5).

The National Council for Financial Inclusion (CNIF) identifies four components which define and shape the condition of financial exclusion: the supply of financial services, the demand for financial services by consumers, institutional protection to financial consumers, and financial education (CNIF, 2016: p 3).

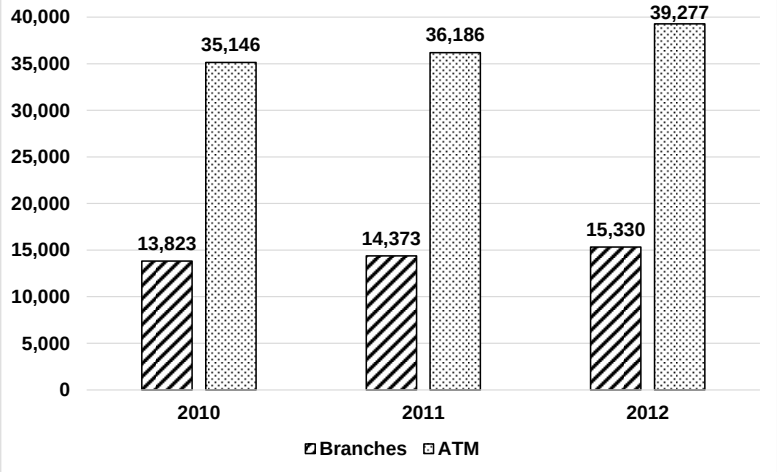
Figure V.1 shows the total number of branches and automatic teller machines (ATM) operating in Mexico between 2010 and 2012, as a proxy of financial supply. The growth rate of branches within this period was 10.9%, whereas ATM grew 11.8%. In 2011, only 65.0% of 2,458 municipalities in the country,

---

52 Over 60% of total workers in Mexico work in the informal sector without labour benefits (ENOE, 2011).

where 96.0% of total adults lived, had at least one financial institution (CNIF, 2012: p 9).

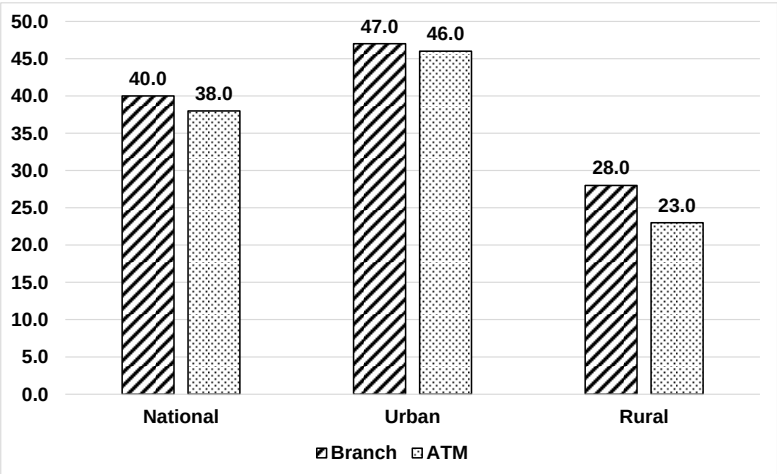
Figure V.1 Total number of branches and ATM in Mexico, 2010-2012



Source: National Council for Financial Inclusion (CNIF, 2016: p 5).

Figure V.2 presents the percentage of adult population using branches or ATM in 2012 at national, urban and rural areas. The use of these services was more extensive in branches at urban areas (47.0%) followed by ATM (46.0%), whereas in rural areas the use of branches was only 28.0%, and ATM was half than the urban case (23.0%).

Figure V.2 Percentage of adult population using branches or ATM, 2012

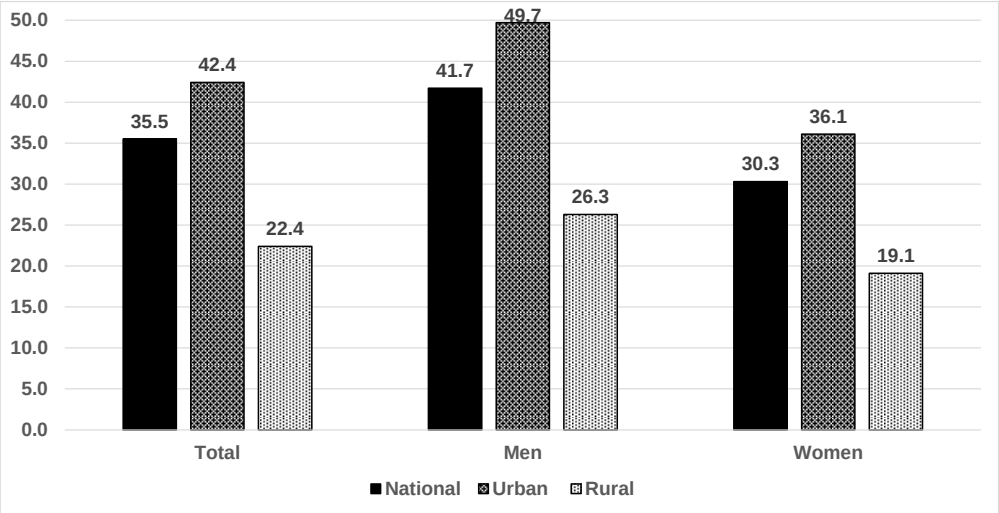


Source: National Council for Financial Inclusion (CNIF, 2016: p 7).

Figure V.3 presents the national percentage of adult population which report ownership of at least one financial instrument, and by urban and rural areas. Nearly 50.0% of men in urban areas report ownership of a financial

instrument. In rural areas, 19.1% of women report ownership of at least one financial instrument.

Figure V.3 Percentage of adults with at least one financial instrument, 2012



Source: National Council for Financial Inclusion (CNIF, 2016: p 8).

The National Commission for the Protection and Defence of Financial Services Users (CONDUSEF) is the institution which protects financial consumers. In 2011, this institution received 1,390,000 complaints, advice and advocacy requests from clients of financial institutions (CNIF, 2012: p 9). Figures on financial education are difficult to follow due to the vast number of institutions involved.

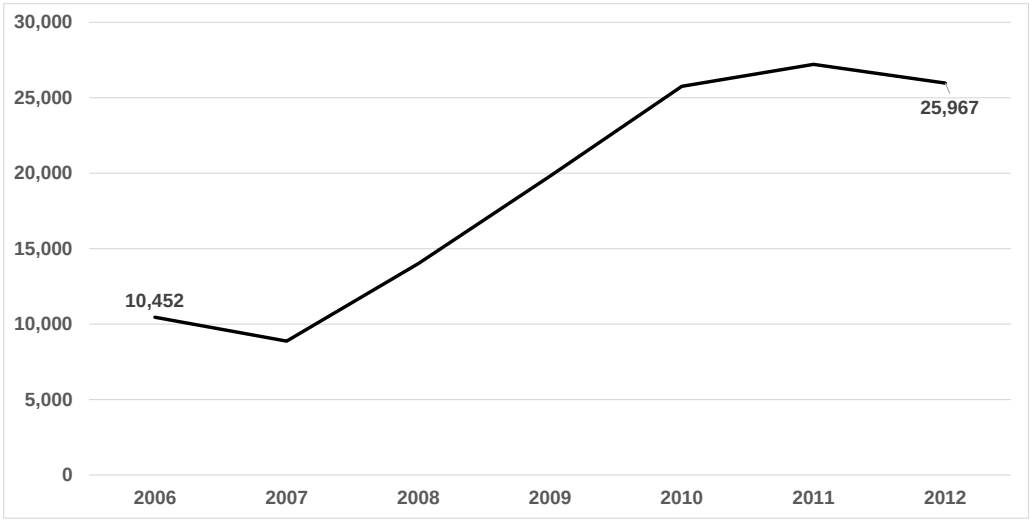
V.2.2 Crime and violence in Mexico

Historically, the border between Mexico and the USA has been a peaceful area since its formal establishment in 1848 (Etter and Lehmuth, 2013). However, the North-American policy of drug prohibition in the early 20th Century generated incentives for illegal drug-smuggling activities at the south of this border (Morales-Oyarvide, 2011). These activities intensified in Latin America during the second half of the 20th Century (Morales-Oyarvide, 2011), due to the increase in drug consumption, the emergence of new international drug cartels, and the weakness of Latin American institutions (Serrano, 2010). As a result, the prohibitionist policy in the USA became more severe since Nixon's administration (Daniel-Rosen and Zepeda-Martinez, 2015).

Drug-traffic was tolerated by Mexican authorities for decades, particularly during the 80 years of the Party of the Institutional Revolution (PRI) rule. When a presidential candidate from the opposition (the National Action Party, PAN) got elected in 2000, the relations between drug cartels changed (Etter and

Lehmuth, 2013). In 2006, a recently inaugurated president (PAN) launched a military campaign against drugs (Daniel-Rosen and Zepeda-Martinez, 2015). This strategy triggered the most devastating wave of violence since the Mexican Revolution in 1910-20 (Correa-Cabrera, 2012). Figure V.4 shows the trend of the total number of homicides in Mexico during this period, which increased 148.8%. These levels of violence had profound effects on society, particularly in the north-eastern States bordering the USA (Correa-Cabrera, 2012).

Figure V.4 Total number of homicides per year, 2006-2012



Source: Information data bank, INEGI, 2013.

V.3 Literature review

This section reviews the literature on the topics of financial exclusion, the role of information in credit markets, social networks, social capital, and elaborates on the link between social networks and the tame of information asymmetries in credit markets. The discussions developed in this section bring a theoretical background on which the empirical analysis of this chapter is supported, as well as the foundations to understand the estimation results.

V.3.1 Financial exclusion

Financial services (credits, savings, insurance) can be classified in two groups: formal and informal. The former are regulated and supervised by governmental authorities, whereas the latter has no supervision (Mansell-Cartens, 1995: p 70). Floro and Ray (1997) identify formal financial institutions

as commercial banks, thrift banks, rural banks, government banks, and non-financial institutions like insurance companies; on the other hand, informal services are transactions between relatives, friends, and unregulated credit cooperatives or credit associations.

Informal financial services have ancient origins; they were developed as a way to canalise and mobilise resources within a group or community (Seibel, 2001). However, in a modern context, informal financial services are usually expensive and charge higher interest rates than the formal ones (Hoff and Stiglitz, 1996). These conditions could lead people into poverty traps (Dymski, 2014).

A generalised definition of financial exclusion is not straightforward due to the different characteristics of the financial markets (Kempson and Whyley, 1999: p 2), and the differences in markets between countries and regions (Hoff and Stiglitz, 1996). Therefore, simple definitions of financial exclusion (without contextualisation) are problematic (Devlin, 2005).

Financial exclusion has been widely discussed in the UK since 1997, as a result of the government's strategy to address this issue (Dymski, 2014). The literature presents several definitions of financial exclusion. For instance, Leyshon and Thrift (1995) define financial exclusion as processes that prevents disadvantaged groups from accessing the financial system. For Kempson and Whyley (1999), financial exclusion is all the constraints in the access to mainstream financial services.

A fundamental approach to understanding financial exclusion is through the broader concept of social exclusion (Sinclair, 2001: p 13), which can be defined as the condition of segregation individuals face from social mainstream activities (Glennerster et al., 1999). Gordon et al. (2000: p 11) define four domains of social exclusion: impoverishment, labour market exclusion, services exclusion, and exclusion from social relations. From this perspective, financial exclusion could be the result of social (services) exclusion in the domain of financial processes.

Some institutions amalgamate the academic definitions into wider ones. The European Commission (EC) defines financial exclusion as a process which hinders the access or use of financial products and services in the mainstream market (EC, 2008: p 9). In Mexico, the National Centre for Financial Inclusion (CNIF) defines financial exclusion as the peoples' lack of access to financial services like payment systems, savings, credit, or insurance (CIF, 2009: p 4).

The causes of financial exclusion could be extensive. Some aspects behind the financial exclusion identified by Kempson and Whyley (1999), Sinclair (2001: p 15), Devlin (2005), and Leyshon and Thrift (1993) are:

1. **Access exclusion**. Financial institutions refuse to provide their services to particular groups of population (e.g. age groups).
2. **Condition exclusion**. Financial institutions set entrance barriers (e.g. minimum deposits, minimum balance, and collaterals).
3. **Price exclusion**. Fees and commissions unaffordable to certain people.
4. **Marketing exclusion**. Promotion, image or the way services are delivered discourage certain people from using financial services.
5. **Self-exclusion**. People do not take or use financial services.
6. **Resource exclusion**. People have an interest in financial services, but not enough resources (e.g. households without enough income for savings).
7. **Location-based exclusion**. A null offer of financial services in a specific geographical area.

Self-exclusion considers the situation when individuals, after an active and unconstrained decision (Kempson and Whyley, 1999), prefer not to use a financial service, despite their need, due to past refusal(s), confusion, negative impressions or lack of trust on financial institutions or financial products (Devlin, 2005).

Non-use of financial services is a broader condition, which includes financial exclusion as well as not needing a particular type of financial product (Devlin, 2005). If the motivations behind the individual's avoidance of financial services are not clear, self-exclusion could be confused with financial exclusion instead of the lack of need for financial products (Sinclair, 2001: p 15).

However, Kempson and Whyley (1999) consider financial exclusion as a dynamic process in nature, defined by intertemporal changes in preferences for the demand for financial services. Then, self-exclusion could also be considered as a condition of exclusion at certain point in time.

V.3.2 Information and credit

One of the main concerns on credit allocation is the lender's lack of complete information about the borrower's repayment intentions (asymmetric information). Not all borrowers have the same degree of risk regarding credit repayment: they could be low or high-risk borrowers (Ray, 1998: p 553). This issue leads to an agency problem, where the lender (principal) has no

knowledge of the borrower's (agent) intentions on repayment, situation which reflects moral hazard[53] (Armendariz and Morduch, 2010: p 41).

In order to differentiate between high/low-risk borrowers, financial institutions have to develop screening devices. Stiglitz and Weiss (1981) propose the use of collateral and interest rates as screening instruments when all borrowers are identical and, therefore, direct observation of the borrowers' characteristics does not solve the information asymmetry.

High interest rates reflect high default rates and high information costs (Stiglitz and Weiss, 1981). Then, interest rates can be used as an indirect instrument to identify the borrowers' willingness to pay for credit. High-risk borrowers are willing to pay higher interest rates than low-risk borrowers when the borrower's returns on the loan are high (Hoff and Stiglitz, 1996).

Collateral (guarantees over loans) are an entry barrier to the credit market ("condition exclusion", as described before). However, if expected returns of risky borrowers are higher compared to low-risk borrowers, then more loans will be assigned to riskier borrowers, even if entry barriers are high (Stiglitz and Weiss, 1981).

The rise of both interest rates and collaterals (increasing entrance barriers) will not reduce risk behaviour from consumers because it depends on the returns of successful credits, not on the increase of barriers itself (Bester, 1985). Therefore, the increase of entrance barriers would exclude borrowers, even low-risk ones, creating an adverse selection condition, and an imperfection in the credit market (Armendariz and Morduch, 2010: p 41).

In informal financial markets, lenders include screening costs in the interest rates they charge (Aleem, 1990; Hoff and Stiglitz, 1996). A partial solution for screening is the peer-monitoring, where groups of clients could monitor each other in order to fulfil their financial obligations (Hoff and Stiglitz, 1996). Several micro-finance schemes are based on peer-monitoring models worldwide (Bateman, 2010: p 7).

---

53 Adverse selection and moral hazard are concepts introduced in the economic analysis by Akerlof (1970) showing how, in an economic transaction, one agent can take advantage from other agent based on the information the first agent has (information asymmetry). Adverse selection is created when, due to this information asymmetry, lenders will reject both low and high-risk borrowers equally (the screening mechanism is not working). Moral hazard is a change in the borrower's risk behaviour due to a new condition, usually product of the adverse selection.

On the other hand, Villa-Boas and Schimdt-Mohr (1999) found that highly differentiated financial products and less competitive financial markets help financial institutions to perform less intensive screening of their clients due to self-selection (clients will allocate themselves into their most appropriated financial product). Gehrig (1998) confirmed this finding. Agarwal et al. (2011) show how soft (non-quantitative) information, collected from the social relations with customers, could enhance the screening and the lending decision processes.

V.3.3 Social relations and social networks

Individuals living and interacting in a society have a set of social relations (links between people) which have a profound influence in almost all areas of their lives (House et al., 1998; Wentzel, 1998; Granovetter, 1985). These social relations, however, are not all the same; they depend on different factors which can help to classify the relations.

For instance, Granovetter (1973) classifies social relations as strong or weak depending on the strength of the peoples' ties (e.g. family, friends, neighbours). Coleman (1988) highlights the relevance of social distance between individuals (e.g. social class, relations of authority) as an essential factor for the strength of those ties and classifies social relations as horizontal structures when individuals are similar, and vertical structures for relations of authority. Putnam (2000: p 21) differentiates social relations according to the level (norms) of reciprocity within the relation.

A set of social relations which face similar characteristics could be defined as a social network, and the rules of conduct within these networks are shaped by mutual obligations and reciprocity (Putnam, 2000: p 20). Peoples' activities can also be embedded at different levels of social relations; the more embedded in social relations activities are, the more socialised these activities can be (Granovetter, 1985).

A vast literature shows the influence of social relations on peoples' lives. In education, social relations can canalise motivation through vertical structures (parents, teachers and peers) in order to achieve academic goals (Wentzel, 1998). Stress (Cutrona & Russell, 1987) and premature deaths (House et al., 1988) are linked to low quantity social relations. Frey (1997: p 26) shows how the strength of the social relation between employer and employee can improve productivity. Granovetter (1973) found how social networks could help people to find a job faster than without using them.

V.3.4 Social capital

From an economic perspective, Bourdieu (1986) defines social capital as the aggregate of actual and potential social networks which could be transformed into economic outputs. Putnam (2000: p 19) defines social capital as a set of social networks which have effects on production and income. Robison and Siles (2012) define social capital as the sympathy one person or group has toward others, which could be capitalised in economic outcomes or preferential treatments. Social capital is a complement to physical and human capital; however, it can produce positive or harmful outcomes (Ostrom, 1999).

Putnam (2000: p 296) argues that trust, networks and reciprocity shape social capital. Durlauf and Fafchamps (2004) consider that the externalities of social capital are shared by its owner(s) due to the relations of trust, norms and the values which shape the network. Social capital could also be determined by the degree of collective action in a community (Ostrom, 2000).

Social capital, as well as social relations, could be classified into different types according to its characteristics. Coleman (1988) identifies three forms of social capital based on the obligations and expectations between the members of the social network, the degree of openness in the information flows within the network, and the set of norms and sanctions which rule the network.

Woolcock (1999) proposed a classification of social capital based on peoples' identities and the strength of the social ties within the network, identifying three categories as follows:

1. **Bonding**, for exclusive networks between individuals with similar identities and strong ties (e.g. family members, friends).
2. **Bridging**, for inclusive networks between individuals with different identities, weak ties, but horizontal social structures (clubs and associations).
3. **Linking**, for vertical social structures with a figure of authority (government, state institutions, justice system).

This classification is widely popular in the literature, and was adopted by international organisations like the United Nations Development Program

(UNDP) and the World Bank as an empirical approach to the measurement of social capital[54].

Solidarity and trust within a social network could be stronger when legal and institutional mechanisms are absent, working as an alternative to formal institutions (Lopez-Rodriguez and De la Torre, 2010). However, violence and crime can deteriorate social capital, destroying the social cohesion within the group (Moser and Holland, 1997) and damaging trust in people, making trade and cooperation more difficult (Colletta and Cullen, 2000). Social networks are not constant: preferences on the use of networks change over time (Sandoval and Lima, 2014).

V.3.5 Social networks and information asymmetries

Social networks could overcome market information asymmetries by means of the trust embedded in the network and expectations of reciprocity, particularly in low- and middle-income countries where formal institutions are not strong (Bloom et al., 2008). Cormier et al. (2009) found evidence that the disclosure of hard (quantitative) information through firms' social networks could reduce market information asymmetries in stock markets. Wu (2008) demonstrates that firms which share information through their social networks increase productivity by reducing market uncertainty.

Lin et al (2009; 2013) found that, in an unsecured (not supervised by official authorities) online peer-to-peer money lending market (P-2-P), borrowers increased their probabilities for credit acceptance and better deals in interest rates when lenders used the borrowers' social networks as a screening mechanism to evaluate their credit trustworthiness.

Lapavitsas (2007) argues that debt repayment could be backed not only by collateral or contracts but also by social and familial sanctions. Moreover, Lapavitsas considers traditional credit information models as narrow because they only consider one side of the trust relation (lender to borrower), when trust also flows in the opposite direction.

In an experimental study in Argentina, Becchetti and Conzo (2011) found that creditworthiness and trustworthiness are closely related (almost synonyms) because, under conditions of asymmetric information, giving a loan highlights

---

54 Some examples of the institutionalisation of this definition can be found in Grootaert et al. (2004), and the Mexican National Social Capital survey (ENCAS) 2011.

trustworthiness in the borrower. Mwangi and Ouma (2012) argue on the relevance of considering existing trust relations within social groups in the allocation of formal credit.

V.4 Methodology

This section discusses the methods to estimate the differences in real interest rates of informal credit between individuals financially and non-financially excluded, by social class (the middle class, the deprived and the affluent). The first subsection discusses different strategies to measure differences between groups. The following subsection concentrates on the estimation of propensity scores. The final three subsections explore estimation methods for effects in a treatment context.

V.4.1 Estimation of differences between groups

A basic approach to observe differences on certain domain between independent and not correlated groups could be a test on the equality of the means (Wilks, 1946; Mehta and Gurland, 1969). This procedure would test whether the means of both groups are the same.

In the case of finding differences between groups, the test provides the magnitude of this difference. However, correlation among the groups (Mehta and Gurland, 1969), and the variance and the covariance in a multivariate context (Wilks, 1946; Morrison, 1973) could lead to a loss of sensitivity in the test as higher standard errors. The estimation of differences between paired observations[55] could help to overcome these effects (Bradley and Terry, 1952).

An alternative approach to measure these differences between groups could be the application of the principles of experimental analysis and causal effects discussed in section IV.4 of the previous chapter, where individuals exposed to certain condition were considered under treatment and individuals not exposed were deemed the control population.

---

55 Observations paired in terms of characteristics which are normal or similarly distributed.

As discussed before, the randomisation of a treatment is a necessary precondition for this approach (Ashenfelter and Card, 1985; Angrist and Pischke, 2009: p 21). However, this condition is not always met because decisions and the individuals' behaviour could be related to the benefits or profits offered by the treatment (Wooldridge 2002: p 907).

A popular approach in the literature to estimate causal treatment effects is the Propensity Score Matching (PSM), which is applied in a wide range of fields of study (Caliendo and Kopeinig, 2008). Introduced by Rosenbaum and Rubin (1983), this method can control the issues raised by Wilks (1946) and Morrison (1973) of covariance in a multidimensional case.

Moreover, Heckman et al. (1997) and Dehejia and Wahba (2002) argue that PSM could address three major sources of bias frequently found in the computation of nonexperimental estimations by adjusting the matching specifications. These sources of bias are:

1. Selection on unobservable characteristics, as defined by Manski (1995: p 4) and elaborated in section IV.4.1.
2. Failure in defining an area where individuals can find a match (common support condition).
3. Failure to weight treatment and control observations in terms of sample size.

A significant disadvantage of this method is the presence of a natural bias in the estimators calculated from a PSM. These estimators are calculated based on a matching of individuals according to their propensity scores, not randomly; therefore, these estimators could be biased towards the propensity scores (List et al., 2003). Another common critique to this method is the arbitrariness in the propensity scores estimation, and the justification of the technique selected for matching the individuals (Caliendo and Kopeinig, 2008).

V.4.2 The propensity score matching

The PSM is a statistical technique which matches individuals with similar propensity scores in order to estimate an average effect to certain treatment as the difference in outcomes between these matched individuals (Greene 2012: p 935). The propensity score is defined as "the conditional probability of assignment to a particular treatment given a vector of observed covariates" (Rosenbaum and Rubin, 1983: p 41). Following Wooldridge (2002: p 911), the propensity score can be expressed as follows:

$$p(x_i) = P(w_i = 1 \mid x_i) \quad \textbf{(V.1)}$$

where $p(x_i)$, stands for the propensity score for individual *i*, *w* stands for the condition of treatment, and *x* represents an independent variable.

The probability of treatment for the individuals could follow a probabilistic (*probit*) specification defined as:

$$P(w_i = 1|x_i) = \alpha + \sum_q \beta_q \Phi_{iq} + \varepsilon_i \quad \textbf{(V.2)}$$

Where w_{it} stands for the condition of treatment (e.g. financial exclusion) with value of 1 (0 for the control condition, non-excluded); α is the constant; ε_{ist} is the error term; and $\sum_q \beta_q \Phi_{iq}$ represents a set of Φ_{iq} control variables with β_q coefficients. The marginal effects could be computed as:

$$MgFx = \frac{\partial y}{\partial x} \quad \textbf{(V.3)}$$

which follows the same definition as equation (IV.4).

The literature presents different algorithms to match individuals according to their propensity scores. One of the most popular is the nearest neighbour, due to its efficiency and good performance in Euclidean spaces (Gutin et al., 2002). A Kernel-based algorithm (and its variants like *Epanechinkov* or *tricube*) has advantages of non-parametric inference; however, it does not necessarily impose restrictions to balance the distribution between groups as nearest neighbour does, which requires the definition of a bandwidth and calibration of settings (Heckman et al., 1998).

Once the propensity score and the matching method are selected, differences can be computed between individuals belonging to the treatment and control groups but with a similar score. In the following subsections, three estimations methods are explored: the average treatment effect, the average treatment on the treated effect, and the average treatment on the no-treated effect.

V.4.3 Average treatment effect

The average treatment effect (ATE) measures the difference in output for a treated observation compared to a non-treated (control) observation. Following Wooldridge (2002: p 911), the output for a control observation *c* would be:

$$E(y_c \mid x_c,\ w_c = 0) = \mu_0(x_c) \quad \textbf{(V.4)}$$

whereas for a treatment τ observation is:

$$E(y_\tau \mid x_\tau,\ w_\tau = 1) = \mu_1(x_\tau) \quad \textbf{(V.5)}$$

Then, under a cross-section context,[56] the ATE estimator is defined as:

$$\tau_{ATE}\,(x) = \; m_1(x_\tau) - m_0(x_c) \quad \textbf{(V.6)}$$

where τ_{ATE} is the difference between the averages of the estimated outputs for treatment ($m_1(x_\tau)$) and control ($m_0(x_c)$) observations, matched accordingly to their propensity scores.

V.4.4 Average treatment on the treated

Note that the ATE estimator tends to be arbitrarily correlated to the treatment/control condition by construction. The average treatment on the treated (ATT) is a useful estimator to control for this effect (Wooldridge, 2002: p 912), and estimates the difference as:

$$\tau_{ATT} = \; E[m_1(x_\tau) - m_0(x_c) \mid w = 1] \quad \textbf{(V.7)}$$

However, because $m_0(x_c)$ is an empty set (all observations have to be treated *w*=1), then they are substituted with treated observation $m_1(x_\tau)$, paired with PSM.

V.4.5 Average treatment on the non-treated (ATNT)

The average treatment on the no-treated (ATNT) follows the same logic as the ATT, but for the opposite case (no-treated observations). This estimator is defined as:

$$\tau_{ATNT} = \; E[m_1(x_\tau) - m_0(x_c) \mid w = 0] \quad \textbf{(V.8)}$$

Then, the empty set $m_1(x_\tau)$ of observations is substituted with $m_0(x_c)$ observations, paired with PSM.

---

[56] Chapter IV calculated an ATE estimator as the DID estimator β_3 in equation (IV.2), and the ATT estimator as equation (IV.5), under a repeated cross-sectional context. Those estimators are different in nature to this cross-sectional case, but similar in their interpretation.

V.5 Empirical analysis

This section presents the results for the empirical analysis of this chapter. The first subsection describes the data used in the estimations. The second subsection presents the middle-class identification strategy. The following subsection shows descriptive statistics on the variables used for the empirical analysis. Finally, the estimation results are presented in the last subsection.

V.5.1 Data

Information on the exact amount households borrow and repay on informal credit, and the lenders' characteristics, is fundamental for this empirical analysis. Few surveys in Mexico collect information on loans. The National Income and Expenditure Household survey (ENIGH) and the Socioeconomic Conditions Module survey (MCS-ENIGH), used in previous chapters, do not contain information on loans or savings.

In 2011, the Mexican Ministry for Social Development (SEDESOL), and the United Nations Development Program (PNUD), collected the National Social Capital Survey (ENCAS). This survey, collected at household level, contains information on household characteristics, assets, and the household's social networks. A particular set of questions about informal credit[57] is also included, which details the amounts borrowed and repaid by the household, and the lender's relationship to the borrower.

Moreover, ENCAS is an open-access survey, statistically representative at national, urban and rural levels[58], but not at State level. Three regions (north, centre and south) group all the 32 Mexican States. Table V.1 presents the distribution of States across the regions.

---

57 Money borrowed from informal non-professional lenders like friends, relatives and co-workers.

58 Rural areas are considered as towns with less than 10,000 inhabitants.

Table V.1 State distribution across regions, ENCAS 2011

Region 1, North	Region 2, Centre	Region 3, South
Baja California	Aguascalientes	Campeche
Baja California Sur	Colima	Chiapas
Coahuila	Distrito Federal	Guerrero
Chihuahua	Guanajuato	Oaxaca
Durango	Hidalgo	Quintana Roo
Nayarit	Jalisco	Tabasco
Nuevo Leon	Estado de Mexico	Veracruz
San Luis Potosi	Michoacán	Yucatan
Sinaloa	Morelos	
Sonora	Puebla	
Tamaulipas	Queretaro	
Zacatecas	Tlaxcala	

Source: ENCAS, 2011.

Considering the alarming crime figures in Mexico, as described in section V.2.3, and following Moser and Holland (1997) as well as Colletta and Cullen (2000), crime should be an important factor to consider in this analysis. Then, the municipal ratio of total homicides related to drug-trafficking in 2011 per 100,000 inhabitants was calculated. The calculation of this index merged data from the Mexican National Census 2010 (Censo, 2010), and the monthly homicide reports published by the office of the Mexican General Attorney (PGR) in 2011. This index would exclude local minor crime but would take into consideration the drug-related crime which has a more profound impact on society.

V.5.2 Middle-class identification

The ENCAS survey is not designed to collect detailed information on household characteristics, income or expenditure, either on labour conditions. Moreover, this survey was not collected by the National Institute of Statistics and Geography (INEGI). The variable construction in ENCAS, codifications and design, follows different methods than INEGI surveys. Then, ENCAS is not compatible with surveys used previously.

This situation represents a challenge regarding middle-class identification. A multidimensional identification, following the methodology discussed in chapter III, is not possible; a unidimensional approach is the most plausible alternative. ENCAS contains information on the total monthly household income. When a respondent did not reveal the exact household income, a table with ranges of income was provided, so she can select the category

which reflects her household income. These income categories were encoded into monetary units using the middle value of the selected range.

Thus, the household per capita income was calculated as:

$$PCI_h = \frac{Total\ Income_h}{HH\ Size_h} \quad \textbf{(V.9)}$$

Where PCI_h represents the per capita income of household h; $Total\ Income_h$ stands for the total income reported in the household h; and $H\ Size_h$ is the household size[59].

From all unidimensional middle-class estimation approaches discussed in chapter II, only absolute and relative definitions allow for its identification. Relative definitions are better for time comparisons (Birdsall, 2010; Ferreira et al., 2013). Considering the cross-sectional nature of the analysis in this chapter, an absolute middle-class definition would be the most appropriated approach.

Table V.2 presents the percentage of population by social class and geographical strata, using different absolute definitions in daily USD PPP[60]. As found in chapter II, these estimations could be divided in two groups: definitions between $2 and $13 USD PPP (Ravallion, 2010; Banerjee and Duflo, 2008), and definitions between $10 and $50 USD PPP (World Bank, 2013; Milanovic and Yitzhaki, 2002). In this case, the former group overestimates the affluent and underestimates the deprived, whereas the latter group is more balanced. This result suggests an income distribution skewed to the right side of the middle-class thresholds[61].

Table V.2 Class estimations using different middle-class definitions, percentages

Middle class definition	National			Urban			Rural		
	Middle class	Deprived	Affluent	Middle class	Deprived	Affluent	Middle class	Deprived	Affluent
$2 ≥ y_i ≤ $13	33.8	3.5	62.7	28.8	2.6	68.6	38.9	4.3	56.8
$10 ≥ y_i ≤ $50	66.7	19.1	14.2	67.7	14.5	17.7	65.6	23.8	10.6
$2 ≥ y_i ≤ $10	15.7	3.5	80.9	11.9	2.6	85.5	19.5	4.3	76.2
$12 ≥ y_i ≤ $50	62.1	23.7	14.2	63.9	18.4	17.7	60.2	29.1	10.6
Total population	**5,391**	**5,391**	**5,391**	**2,722**	**2,722**	**2,722**	**2,669**	**2,669**	**2,669**

Source: own estimations using microdata from ENCAS, 2011.

59 Descriptive statistics of the daily income distribution are presented in Appendix D.1.

60 The $\mathrm{PCI_h}$ was divided by 30.4 to obtain values in daily terms.

61 Figures for the income density at national, urban and rural contexts are available in Appendix D.2.

Considering these results, the definition of middle class in this chapter is the one with thresholds between $10 and $50 USD PPP, because has the most balanced identification of the deprived and the affluent. Besides, this definition has the best matching between multidimensional and unidimensional methods as found in chapter III.

V.5.3 Descriptive statistics

The ENCAS survey covers 5,391 households nationwide, 2,722 in urban areas and 2,669 in rural areas. Households are evenly distributed across regions (1,796). The survey considers only one respondent per household with a minimum age of 17 years, regardless gender, age or household position (e.g. household head). All data information at individual level corresponds to the respondent.

This empirical analysis focuses exclusively on urban areas in order to control for location-based exclusion bias. Moreover, the sample is restricted to all individuals who had an informal loan in the last 12 months before the interview (498 observations). Then, the sample is divided into three groups for each social class.

Financial exclusion is defined as the absence of formal financial instruments in the household. This definition could be arguable because it could refer to non-use of financial services instead of exclusion (Devlin, 2005; Sinclair, 2001: p 15). Unfortunately, ENCAS has no information about the motivation behind not having financial services. Then, the definition in this work follows Kempson and Whyley (1999), considering self-exclusion as a type of financial exclusion at certain moment in time. Information on trust on institutions (linking-type networks) is included in the estimations in order to control for the lack of trust in institutions as a reason behind financial exclusion (Devlin, 2005).

Informal credits are loans given only by neighbours, friends, *compadre/comadre*[62], relatives or co-workers to the borrower (respondent) in

---

62 *Compadre* (and its female *Comadre*) is a social figure between parents and godparents formed on the occasion of Christian rituals (baptism, first communion, confirmation). This link is “semi-sacred”, usually stronger than family links (e.g. cousins), and it is often compared to ties between siblings (Friedrich, 1965).

the last year. Only the loan with the largest amount is taken into consideration. The interest rate was calculated as follows:

$$RI_c = \frac{Amount\ paid_c}{Amount\ borrowed_c} \quad \textbf{(V.10)}$$

Where RI_c stands for the real interest rate on loan *c*; $Amount\ paid_c$ is the amount paid for loan *c*; and $Amount\ borrowed_c$ is the amount of money borrowed for loan *c*.

A total household dependency ratio (considering children and the elderly) was estimated in order to control for sociodemographic dynamics within households. An asset index was also estimated in order to control for household wealth instead of income because assets could work as collateral. Details on the estimation of the assets index are available in Appendix D.3.

ENCAS information on social networks measures the level of trust individuals have in three types of people (family, friends, and *compadre/comadre*), and in three State institutions (the police, the government, and judges). Respondents were asked to score their trust in each of these six figures in a 0-10 scale (0 for no trust, 10 for total trust). Then, trust in social networks is clustered into bonding-type networks (family, friends, *compadre/comadre*), and linking-type networks (the police, the government, and judges). Thus, standardised indices for trust in bonding and linking networks were estimated as follows:

$$SCI_{ik} = \left(\frac{\sum_{i=1}^{t} S_{ikt}}{3}\right)/10 \quad \textbf{(V.11)}$$

Where SCI_{ik} is the average trust score of individual *i* on the *k*-type social network (bonding or linking); and $\sum_{i=1}^{s} S_{isk}$ stands for the sum of all *t* trust scores of individual *i* in the *k* social network. Finally, SCI_{ik} is standardised in order to have values within 0 to 1.

Table V.3 presents descriptive statistics for the individuals in the middle class. The middle class has 371 observations, from which only 360 provided complete information to calculate real interest rates (RI_c). The real interest rate was, on average, 10.2%. 78.4% of these individuals report being financially excluded. 41.8% are men. The average age is 36.7 years, and the average level of education in years is 10.5. 45.6% are household heads, and 60.6% are working. The average household dependency ratio is 0.650.

Table V.3 Descriptive statistics, the middle class

Variable	Observations	Mean	Standard deviation	Minimum	Maximum
Real interest rate*	360	10.2	70.2	0.0	900.0
No formal finance*	371	78.4	41.2	0.0	100.0
Male*	371	41.8	49.4	0.0	100.0
Age	371	36.7	13.6	18.0	82.0
Age squared	371	1,532.9	1,173.7	324.0	6,724.0
Years of education	371	10.5	4.7	0.0	19.0
HH head*	371	45.6	49.9	0.0	100.0
Working*	371	60.6	48.9	0.0	100.0
HH depndency rate	371	0.650	0.653	0.000	3.000
HH assets index	371	0.777	0.129	0.344	0.983
Bonding networks	371	0.765	0.171	0.000	1.000
Linking networks	371	0.537	0.244	0.000	1.000
Homicide rate[1]	371	9.40	17.59	0.00	114.05
Region 1*	371	24.8	43.2	0.0	100.0
Region 2*	371	35.6	47.9	0.0	100.0
Region 3*	371	39.6	49.0	0.0	100.0
Amount borrowed	366	2,998.3	6,271.7	50.0	100,000.0

Source: own estimations using microdata from ENCAS, 2011.
[1] Ratio of yearly drug-trafficking related homicides in the municipality per 100,000 inhabitants.

Regarding trust, the average score on bonding (people) networks (0.765) is higher than on linking (institution) networks (0.537). The yearly homicide rate is 9.4 murders per 100,000 inhabitants. Individuals are relatively concentrated in regions centre and south, whereas the north has the lowest incidence (24.8%). The average amount of informal credit is $2,998.30 Mexican pesos.

Table V.4 presents descriptive statistics for the individuals in the deprived class. This group has 67 observations, 64 with complete information in real interest rates. The average real interest rate was 7.9%, 86.6 of the sample is financially excluded, and 40.3% are men. This group is older on average than the middle class (38.7 years) and less educated (8.4 years). Almost 60.0% are household heads, and half of them are working. The total household dependency ratio is lower than in the middle class (0.499), as well as the household assets index (0.679).

Table V.4 Descriptive statistics, the deprived

Variable	Observations	Mean	Standard deviation	Minimum	Maximum
Real interest rate*	64	7.9	24.5	0.0	135.3
No formal finance*	67	86.6	34.4	0.0	100.0
Male*	67	40.3	49.4	0.0	100.0
Age	67	38.7	14.7	18.0	72.0
Age squared	67	1,713.6	1,277.7	324.0	5,184.0
Years of education	67	8.4	5.8	0.0	17.0
HH head*	67	59.7	49.4	0.0	100.0
Working*	66	50.0	50.4	0.0	100.0
HH depndency rate	67	0.499	0.638	0.000	3.000
HH assets index	67	0.679	0.136	0.344	0.983
Bonding networks	67	0.756	0.176	0.267	1.000
Linking networks	67	0.480	0.259	0.000	1.000
Homicide rate[1]	67	2.43	6.95	0.00	33.93
Region 1*	67	6.0	23.9	0.0	100.0
Region 2*	67	34.3	47.8	0.0	100.0
Region 3*	67	59.7	49.4	0.0	100.0
Amount borrowed	64	3,271.6	8,985.2	50.0	50,000.0

Source: own estimations using microdata from ENCAS, 2011.
[1] Ratio of yearly drug-trafficking related homicides in the municipality per 100,000 inhabitants.

Deprived population have similar average score of trust in bonding networks as the middle class (0.756), but lower trust in institutions (0.480). They live in municipalities with lower exposition to crime (2.43), are concentrated in the south (almost 60.0%), and have almost no presence in the north (6.0%). The average of informal credit is higher than in the middle class ($3,271.60 pesos).

Finally, Table V.5 shows descriptive statistics for the affluent class. This group has 60 observations in the sample, 55 with complete information. They pay higher real interest rates (38.8%). This result could be explained because they also have the highest credit amounts ($5,560.70 pesos). 76.7% report no ownership of formal finance instruments (an unexpected outcome considering their social position and the average of 50% in the urban context, although lower than the middle and deprived). 53.3% are male, have an average age of 38.2 years, and have the highest average on education (11.7 years). Very few of them are household heads (33.3%), and 65% report having a job. Their dependency ratio is the lowest (0.384), and this group has the highest average on the assets index (0.867).

Table V.5 Descriptive statistics, the affluent

Variable	Observations	Mean	Standard deviation	Minimum	Maximum
Real interest rate*	55	38.8	169.9	0.0	900.0
No formal finance*	60	76.7	42.7	0.0	100.0
Male*	60	53.3	50.3	0.0	100.0
Age	60	38.2	14.5	18.0	72.0
Age squared	60	1,667.4	1,249.0	324.0	5,184.0
Years of education	60	11.7	4.4	0.0	20.0
HH head*	60	33.3	47.5	0.0	100.0
Working*	60	65.0	48.1	0.0	100.0
HH depndency rate	60	0.384	0.536	0.000	2.500
HH assets index	60	0.867	0.117	0.509	0.983
Bonding networks	60	0.770	0.160	0.333	1.000
Linking networks	60	0.462	0.244	0.000	1.000
Homicide rate[1]	60	16.18	32.48	0.00	205.54
Region 1*	60	41.7	49.7	0.0	100.0
Region 2*	60	35.0	48.1	0.0	100.0
Region 3*	60	23.3	42.7	0.0	100.0
Amount borrowed	56	5,560.7	10,198.2	200.0	70,000.0

Source: own estimations using microdata from ENCAS, 2011.
[1] Ratio of yearly drug-trafficking related homicides in the municipality per 100,000 inhabitants.

This group has a level of trust in bonding networks of 0.770, just between the deprived and the middle class, but the lowest trust in institutions (0.462). They also live in the most conflictive regions (16.18 on the annual homicide rate) and are concentrated in the north and the centre regions (41.7% and 35.0%).

A possible explanation on the low levels of crime and violence faced by the deprived classes, and high levels faced by the affluent, could be the geographical distribution of the observations. At this time (2011), the north of Mexico was facing the worst effects of the government strategy against the drug-trafficking, particularly on the USA border.

V.5.4 Estimation results

Following a quasi-experimental framework, individuals identified as financially excluded are defined as treatment observations, whereas individuals with formal financial products are defined as control. In this way, the treatment effects will be estimated on the condition of being excluded from the formal financial sector.

As mentioned in section V.4.1, a test on the equality of the means between the treatment and control groups was conducted. Table V.6 presents the differences on the equality of the means in real interest rates RI_c between the treatment and control groups. These differences are computed for the social classes and the overall urban context.

Table V.6 Test on the equality of means, real interest rates

Group	Urban context	Middle class	Deprived	Affluent
Control	0.0599	0.0445	0.0512	0.1659
	(0.0187)	(0.0172)	(0.0292)	(0.1037)
Treatment	0.1509	0.1184	0.0840	0.4501
	(0.0483)	(0.0470)	(0.0353)	(0.2917)
Difference	-0.0911	-0.0738	-0.0328	-0.2842
	(0.0953)	(0.0899)	(0.0886)	(0.5585)
Pr(|T| > |t|)[A]	0.3396	0.4118	0.7130	0.6130

Source: own estimations using microdata from ENCAS, 2011.
[A] p-value associated with the equality of the mean across gender.
*Note: Standard errors in brackets.

Despite negative differences in all groups, meaning higher interest rates on the financially excluded compared to the non-financially excluded, these differences are not statistically significant at any conventional level of significance. Therefore, this analysis shows no differences in real interest rates for any group.

These results could be explained due the limitations of the analysis on means, like the lack of control for the correlation among the groups (Mehta and Gurland, 1969), and the variance and the covariance of the determinants of these conditions (Wilks, 1946; Morrison, 1973), a possible loss of sensitivity (high standard errors) could explain these non-significant results.

Considering these concerns, the PSM estimation is suggested. Following Caliendo and Kopeinig (2008), a first step on a PSM estimation is the definition of a set of independent variables to compute the propensity score. A series of probabilistic (*probit*) estimations where calculated on the sample of all individuals living in the urban context[63], explaining financial exclusion following equation (V.2).

Altunbas et al. (2010) explore the determinants of formal financial exclusion in Bolivia, identifying that personal characteristics like belonging to indigenous groups, education, labour condition, and gender are important determinants. Then, following Altunbas et al., a basic model specification has individual and household characteristics (column 1). Then, extra-household information like trust in social networks, crime, and controls by region are included in a second estimation stage (column 2). The average marginal effects of these

63 This sample is not restricted to the condition of having an informal loan in order to explore the covariance between having informal credits to the lack of formal financial services, as warned by Wilks (1946) and Morrison (1973).

estimations, following equation (V.3), are presented in Table V.7 by social class.

Table V.7 Determinants of the lack of formal financial services

Variables	Urban context		Middle class		Deprived		Affluent	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Male	0.0244	0.0238	0.0244	0.0238	0.0180	0.0170	0.0295	0.0293
	(0.0177)	(0.0173)	(0.0177)	(0.0173)	(0.0130)	(0.0123)	(0.0214)	(0.0213)
Age	-0.0058**	-0.0059**	-0.0058**	-0.0059**	-0.0044**	-0.0043**	-0.0071**	-0.0072**
	(0.0025)	(0.0025)	(0.0025)	(0.0025)	(0.0019)	(0.0018)	(0.0030)	(0.0030)
Age square	0.0000**	0.0000**	0.0000**	0.0000**	0.0000**	0.0000**	0.0000**	0.0000**
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Years of education	-0.0032**	-0.0037**	-0.0032**	-0.0037**	-0.0024**	-0.0027**	-0.0039**	-0.0045**
	(0.0015)	(0.0015)	(0.0015)	(0.0015)	(0.0011)	(0.0011)	(0.0019)	(0.0019)
HH head	-0.0302	-0.0292	-0.0301	-0.0292	-0.0225	-0.0212	-0.0366	-0.0361
	(0.0189)	(0.0186)	(0.0189)	(0.0185)	(0.0142)	(0.0135)	(0.023)	(0.0229)
Working	-0.0487***	-0.0520***	-0.0484***	-0.0518***	-0.0379***	-0.0395***	-0.0585***	-0.0635***
	(0.0163)	(0.0158)	(0.0161)	(0.0157)	(0.0130)	(0.0124)	(0.0195)	(0.0192)
HH total	0.0406***	0.0354***	0.0406***	0.0354***	0.0305***	0.0258***	0.0489***	0.0434***
dependency rate	(0.0127)	(0.0123)	(0.0127)	(0.0123)	(0.0096)	(0.0091)	(0.0154)	(0.0153)
HH asset index	-0.4878***	-0.5765***	-0.4877***	-0.5770***	-0.3668***	-0.4200***	-0.5881***	-0.7069***
	(0.0567)	(0.0580)	(0.0566)	(0.0581)	(0.0355)	(0.0354)	(0.0767)	(0.0806)
Informal credit	-0.0721***	-0.0674***	-0.0717***	-0.0672***	-0.0557***	-0.0506***	-0.0862***	-0.0820***
	(0.0206)	(0.0202)	(0.0205)	(0.0201)	(0.0163)	(0.0155)	(0.0244)	(0.0243)
Bonding networks	-	-0.0752	-	-0.0753	-	-0.0548	-	-0.0923
		(0.0532)		(0.0531)		(0.0386)		(0.0656)
Linking networks	-	-0.0809**	-	-0.0810**	-	-0.0589**	-	-0.0992**
		(0.0329)		(0.0330)		(0.0242)		(0.0401)
Homicide rate	-	0.0011***	-	0.0011***	-	0.0008***	-	0.0013***
		(0.0002)		(0.0002)		(0.0002)		(0.0003)
Region 1	-	0.0642***	-	0.0639***	-	0.0448***	-	0.0821***
		(0.0164)		(0.0163)		(0.0112)		(0.0216)
Region 2	-	0.0837***	-	0.0839***	-	0.0590***	-	0.1043***
		(0.0154)		(0.0155)		(0.0110)		(0.0195)
N	2,709	2,709	1,839	1,839	392	392	478	478
Pseudo R^2	0.0603	0.0928	0.0603	0.0928	0.0603	0.0928	0.0603	0.0928

Source: own estimations using microdata from ENCAS, 2011.
Note: Standard errors in brackets; levels of significance at 90%, 95% and 99% as *, ** and *** respectively.

These results suggest most of the variables included in the model explain financial credit (at 95% significance level), except gender (male), being a household head, and trust in bonding-type social networks. On average, the increase of age and education on one year reduces the probability of financial exclusion in 0.5% and 0.4% respectively, suggesting more mature and educated people have access to (or seek for) formal financial services. Having a job also reduces the probability of financial exclusion, perhaps as a result of having a stable source of income.

A positive relationship is found between the total dependency ratio and financial exclusion, possibly related to financial constraints in households with a significant number of economic dependents. The household asset index has a negative relationship to financial exclusion, explained as the availability of collateral in the household, or as an accumulation of capital due to access to credit. Trust in institutions has a negative effect on financial exclusion of 7.0% on average. A rise in homicides has a positive effect of 0.1% on financial

exclusion. Living in the northern and central regions has a positive effect of 6.0% and 8.0% respectively compared to people living in the south.

Following these results, the propensity scores were computed using all the variables included in the second estimation stage plus information on the size of the loan so the credit size will control the matching. These propensity scores are the individuals' probability of not having formal financial products (the probability of being financially excluded). Four propensity scores were estimated, one for each social class and one for the overall urban context. The estimation of these propensities is reported in Appendix D.4.

Finally, the matching of individuals followed the nearest neighbour algorithm due to its simplicity in calculation and less requirement of control specifications as Kernel algorithms. Three difference estimators were computed for the variable real interest rates (RI_c): ATE, (equation IV.6); ATT (equation IV.7); and ATNT (equation IV.8). All standard errors were calculated using bootstrapping (100 replications). Results for these estimators are shown in Table V.8.

Table V.8 Differences in real interest rates by social class

Estimation	Urban context	Middle class	Deprived	Affluent
ATE	0.1087**	0.0734*	0.0356	0.3793
	(0.0550)	(0.0382)	(0.0499)	(0.4175)
ATT	0.1084**	0.0819*	0.0535*	0.4073
	(0.0517)	(0.0429)	(0.0294)	(0.3927)
ATNT	0.1098	0.0426*	-0.0476*	0.2740
	(0.0800)	(0.0323)	(0.0235)	(0.3764)

Source: own estimations using microdata from ENCAS, 2011.
Note: Standard errors in brackets; levels of significance at 10%, 5% and 1% as *, ** and *** respectively.

On average, an urban middle-class individual financially excluded would pay 7.3 percent points more on her real interest rate than an individual who is non-excluded (ATE). Between middle-class financially excluded individuals, the difference of the interest rate to pay for informal loans is 8.2 percent points, whereas for financially included individuals this difference is 4.3 percent points. All these coefficients are statistically significant at 10% level.

Deprived individuals financially excluded have an average difference of 5.3 percent points on their interest rates; a deprived non-financially excluded person have an average of -4.8 percent points. The difference between individuals excluded and non-excluded is not statistically significant, meaning there is no difference in interest rates between them.

Differences in the urban context are statistically significant between the non-financially and financially excluded, as well as within those financially excluded, which represent a difference on interest rates of 10.85 percent points on average for both differences. The ATNT estimator is not statistically significant. Finally, no statistically significant differences are found in the affluent class.

V.6 Conclusion

The objective of this chapter is to analyse differences in real interest rates on informal loans between financially excluded and non-excluded Mexican middle-class individuals. This difference could be interpreted as a real cost of financial exclusion. Considering that this analysis focuses only on loans between closely related people, information asymmetries should be reduced. Then, no difference in real interest rates between excluded and non-excluded individuals should be expected.

The estimation is based on a propensity score matching, which pairs individuals by their probability of not having formal financial products (propensity score), controlled by the amount borrowed. Four propensity scores were estimated, one for each social class and one for the overall population. The analysis considers only urban areas to avoid location-based exclusion bias (Leyshon and Thrift, 1993), and controls by the recent increase in crime experienced in Mexico, which could hurt social trust, trade and cooperation (Moser and Holland, 1997; Colletta and Cullen, 2000).

Results show, opposite to the original expectations, a positive difference on the real interest rate for informal credit of 7.3 percent points between financially excluded and non-excluded middle-class individuals. Moreover, the financially excluded pay 8.2 percent points more on average for informal credit, whereas the included pay only 4.3 percent points more.

In the deprived class, no differences between the financially excluded and non-excluded were found, although the excluded pay 5.3 percent points higher interest rates when comparing among them, and the non-excluded pay less interest rate at 4.8 percent points. No differences at any level were found in the affluent class.

These results could seem counter-intuitive at first glance, because the shorter social distance between borrower and lender represents more trust and

reciprocity (Putnam, 2000: p 20), and should reduce interest rates in loans (Lin et al., 2009; 2013). A possible explanation to this phenomenon is the increase in crime and violence experienced in Mexico since 2006, which could erode the trust in social networks and make transactions more difficult (Moser and Holland, 1997; Colletta and Cullen, 2000).

A significant limitation of this work is the definition of financial exclusion, which may be seen as closer to the non-use of formal financial services rather than to financial exclusion proper (Devlin, 2005; Sinclair 2001: p 15). This definition could lead to an overestimation of the real interest rate differential by considering individuals with access to formal financial services but who decided not to have these services as excluded.

Unfortunately, there is no information available in Mexico to control by the motivation behind not using financial services which, as mentioned before, would lead to a more accurate identification of financial exclusion. This lack of information is the main driver behind the decision to use non-use of formal financial services as a proxy for financial exclusion in this chapter.

Further research could be developed along these lines. Moreover, more controls by type of crime, not just homicides, could be incorporated. Another important limitation is the sample, which was not designed to identify informal loans but social capital, and therefore could be not accurate for this estimation when the sample is divided in smaller population groups.

Chapter VI
Conclusions

VI.1 Introduction

This thesis attempted to provide a more comprehensive understanding of the middle class in Mexico, from a unidimensional income perspective within the Latin American context, to a multidimensional analysis of the Mexican middle class. Moreover, this work tried to understand middle-class women's labour conditions when facing the introduction of no-fault divorce, as well as to explore the real cost for informal loans faced by middle-class individuals when they are financially excluded.

Traditional income definitions of the middle class are subjective; they depend on arbitrary thresholds. Thus, estimations of the middle-class are sensitive to the threshold selection. Estimations in chapter II showed a consistent expansion in the Latin American middle class in most income definitions during the period 2000 and 2015 (absolute low-threshold definitions suggest a contraction). The bipolarisation analysis confirmed this finding, suggesting a decrease in income polarisation, and inferring an expansion of the middle class within this period.

Unidimensional definitions, however, do not reflect the multidimensional nature of the middle class, as discussed in chapter III. Estimations in this chapter showed no perfect matching (overlap) in the identification of the middle class between unidimensional definitions and the multidimensional method proposed. The best matching observed in the middle-class identification between a unidimensional income definition and the multidimensional counting approach was the $10 and $50 USD PPP absolute definition (34.0% overlap).

Regarding middle-class women, the introduction of no-fault divorce in Mexico City had an expected positive effect on middle-class women labour participation of 3.5 percent points, a decrease of labour participation in the informal sector of 3.1 percent points, and an increase of 1 working hour per week. Financial exclusion, on the other hand, increased the real cost of informal loans on 7.3 percent points among middle-class individuals, an unexpected outcome considering the literature, whereas no statistically significant differences were found for the other social classes. The following sections present a detailed description of the findings of this thesis for each of the empirical chapters.

VI.2 The middle class in Latin America

Chapter II discussed different definitions of, and approaches to, the middle class in Latin America, and presented an empirical analysis of its evolution based on a unidimensional approach, using per-capita household income as wellbeing domain, during the period 2000-2015, for Brazil, Chile, Mexico, Paraguay, Peru, and Uruguay. The middle class was identified based on traditional relative and absolute definitions, as well as bipolarisation analysis.

An appropriate definition of the middle class in Latin America is fundamental, considering the high level of inequality in the region (Daza and Cortes, 2013). However, the lack of consensus on the concept, and the arbitrariness on the thresholds selection, could lead to inconclusive results on the estimation of the middle class. For instance, absolute income definitions designed for developing countries, such as $2 to $13 and $2 to $10 USD PPP per day (Birdsall, 2010), showed a trend opposite (inconsistent) to those observed with the rest of definitions (absolute, relative and bipolarisation).

The bipolarisation analysis showed a consistent decrease in bipolarity during this period for all countries, inferring an expansion of the middle class. However, a discrepancy in estimations between two bipolarisation indices was found in the case of Paraguay, where it was not possible to conclude whether income bipolarisation has increased or decreased. The RBL curve analysis presented a few year-versus-year dominance relationships in all countries (except for Chile and Mexico). These relationships mostly covered mid-term changes in bipolarisation (except for Uruguay where the relation covers only one year gap), and reflected a decrease in bipolarisation over time. The analysis of B areas under the RBL curve as bipolarisation indices supported these results.

Regarding the methodological contributions of this chapter, the advantages of the analysis of RBL areas is the absence of controversial threshold definitions. However, one of its main drawbacks is the absence of direct identification of the individuals who belong to the middle class. Another current limitation of the bipolarisation analysis is the absence of an analytical method to test the statistical differences between distributions. However, statistical inference tools can be developed eventually, drawing on existing methods for inequality measurement.

In conclusion, most unidimensional income approaches showed the middle class was expanding during the period 2000-2015. However, not all definitions were consistent (definitions with lower thresholds showed a contraction of the

middle class). In most cases, traditional middle-class unidimensional definitions were consistent with bipolarisation analysis, except for some periods in Paraguay. Then, a reduction of social polarisation in the region was observed, in opposition to the sociological paradigm of the absence of a middle class in Latin America (Whiteford, 2013).

VI.3 The middle class in Mexico

Chapter III analysed the middle class as a multidimensional construct, and proposed two methods for the identification and measurement of the middle class in Mexico based on the counting approach. The first method consists of selecting a specific set of household, non-income, characteristics which define social classes (deprived, middle, and affluent), counting these characteristics, and identifying to which social class the individuals belong. The second method is a particular case of the counting approach, which combines the indirect identification of the deprived (multidimensional poverty) with the unidimensional middle-class income threshold for the affluent.

Using the sample of the Mexican survey MCS-ENIGH for 2012, the direct estimation of the middle class identified 59.2% of individuals in the sample as middle class, whereas 58.2% and 7.5% were identified as deprived and affluent respectively. Following the analysis proposed by Castellani et al. (2015), intersections between these groups showed that 31.9% of the middle class were lower middle class, 10.7% were upper middle class, and the remaining 57.4% were a mid-middle class.

When comparing the estimations of the counting approach to the absolute and relative middle-class income definitions, the absolute $10 - $50 USD PPP per day definition matched the best with the counting approach (34.0%). Thus, this absolute definition identifies better the multidimensional definition of the middle class proposed in this work than any other income definition. On the other hand, the poorest matching with the counting approach was from the $2 - $10 USD PPP definition, suggesting that low-income thresholds do not reflect the real conditions of a middle-income country like Mexico (Birdsall, 2010), and supported the findings in chapter II (middle class defined with low thresholds had opposite trends to all other estimations).

The second counting approach method proposed, based on multidimensional poverty identification and the upper-middle-class income threshold, showed a high sensitivity to changes in poverty thresholds. The highest middle-class

incidence was found using $50 USD PPP per day (51.1%), and the lowest using $10 USD PPP (34.5%). Unfortunately, the characteristics of the multidimensional poverty definition in Mexico do not allow for comparisons between this method and the first method discussed previously.

One of the main challenges faced when applying the counting approach was the absence of enough and complete information for the definition of *y*-type categorical variables (i.e. those that can help discriminate between the affluent and the non-affluent). This issue came as a result of political prioritisation of poverty alleviation. Consequently, surveys are designed to collect information on vulnerable social groups rather than on the top-end of society.

VI.4 No-fault divorce and female labour market participation: evidence from the Mexican middle class

Chapter IV measured the impact of the introduction of no-fault divorce in Mexico on the labour market participation of women belonging to the middle class, focusing on three dimensions: overall labour participation, informal sector participation and the number of hours worked per week. The definition of the middle class was based on an absolute definition, using a predicted labour income as a wellbeing domain.

Results showed that, on average, the probability of middle-class women participation in the labour market increased by 3.5 percent points, whereas that of participating in the informal sector decreased by 3.1 percent points. Middle-class women worked, on average, one extra hour per week in the aftermath of the legislation's introduction. These results seemed to follow the original motivation of introducing no-fault divorce by the National Human Rights Commission and the Supreme High Court of enhancing women's autonomy and freedoms (CNDH, 2013; SCJN, 2012). However, these results could also reflect Parkman's arguments that a weak institutional protection for women (e.g. narrow definition of property rights, limited alimony) could promote women labour participation when introducing no-fault divorce (Parkman, 1992).

Concerning the estimation strategies, DID estimations based on synthetic panels, the proposed method in this chapter, yielded similar coefficients along with almost the similar significance level in comparison to those obtained by no-panel repeated cross-sections. Synthetic panels have a natural control of fixed effects and less variance within the variables, making them more

efficient. Kernel-adjusted DID estimators overestimated the social intervention effect, particularly in the case of the continuous dependent variable (natural logarithm of working hours). A possible explanation could be that synthetic panels are measured regarding the cohorts' means; therefore, the influence of selection-on-observables on the conditional expectation of the treatment effect is being minimised.

A significant limitation of this analysis was the time frame. Similar studies in developed countries like Jacob (1989), Parkman (1992) and Genadek et al. (2007) use census data, covering a timeframe of decades and thereby suggesting long-term effects, whereas the analysis in this chapter relied on two three-year periods before and after the legal innovation.

VI.5 The real cost of financial exclusion among the Mexican middle class

Chapter V analysed the difference in real interest rates paid for informal loans between financially excluded urban middle-class individuals and those not excluded. The empirical analysis focused on middle-class individuals, but also provided results for those classified as urban deprived and urban affluent for comparative purposes. The differences were calculated for real interest rates on informal loans between financially excluded and non-excluded Mexican middle-class individuals. This difference could be interpreted as a real cost of financial exclusion. Considering that this analysis focuses only on loans between closely related people, information asymmetries should be reduced, and no difference in real interest rates between excluded and non-excluded individuals should be expected.

Opposite to the original expectations, the results showed a positive difference in the real interest rate for informal credits of 7.3 percent points between financially excluded and non-excluded middle-class individuals. In the case of the deprived and the affluent, the difference in interest rates between excluded and non-excluded individuals was in general statistically insignificant.

These results could seem counter-intuitive at first glance because the shorter social distance between borrower and lender represents, more trust and reciprocity would be expected (Putnam, 2000: p 20), and should reduce interest rates in loans (Lin et al., 2009; 2013). A possible explanation to this phenomenon is the increase in crime and violence experienced in Mexico since 2006, which could erode the trust in social networks and make

transactions more difficult (Moser and Holland, 1997; Colletta and Cullen, 2000).

VI.6 Future research agenda

Future research on bipolarisation could be on the interpretation and analysis of the other RBL curve areas calculated in the empirical analysis of chapter 1 (A, C, D), as well as the development of a statistical test of differences which could confirm whether the changes in bipolarisation are statistically significant or not.

Regarding the multidimensional analysis of the middle class, a future research agenda includes considering the regional analysis of the middle class. As observed, the distribution of certain household assets does not follow a normal homogeneous distribution across the country; instead, assets are segmented according to cultural and environmental conditions. This correlation could be used as selection criteria for the definition of a regional middle class.

Another topic for future research in the field of multidimensional analysis is the use of the international multidimensional poverty index (MPI) dataset, which samples multidimensional poverty in a wide range of countries, to identify the deprived class as developed in the second method. In this way, an international comparison of the middle class, with consistent class definitions, could be achieved.

In the topic of difference-in-difference (DID) estimations, future research could explore other repeated cross-section techniques and compare the effectiveness and efficiency of DID estimations based on synthetic panel data. Propensity score matchings could be an alternative with different settings of the nearest neighbour or kernel matchings (radius matching as an alternative to the Kernel matching for instance), as well as exploring different ways to generate the population cohorts.

Finally, future research on financial exclusion could work on the development of the motivations behind it, which could be an important limitation of the current work because it is closely related to the non-use of formal financial services (Devlin, 2005; Sinclair 2001: p 15). Unfortunately, the dataset has no information to control for the motivation behind not using financial services. Further research could develop along these lines. Moreover, more controls for type of crime, not just homicides, could be incorporated.

Bibliography

Abeles, M. Amarante, V. and Vega, D. (2014) “Participación del ingreso laboral en el ingreso total en América Latina, 1990-2010” *Revista Cepal*, No. 114, Chile, p 31-52.

Acemoglu, D. and Robinson, J. A. (2009) *Economic Origins of Dictatorship and Democracy*, Cambridge University Press, US, p 256-258.

Agarwal, S. Ambrose, B. W. Chomsisengphet, S. Liu, C. (2011) “The role of soft information in a dynamic contract setting: Evidence from the home equity credit market” *Journal of Money, Credit and Banking*, Vol. 43, No. 4, p 633-655.

Aguillon-Leon, I. Tinoco-Tapia, G. E. and Vargas-Manzano, E. A. (2010) “Consecuencias económicas y psicológicas del divorcio en las mujeres. Un enfoque micro regional en Pachuca, Hidalgo” *Revista de Investigación en Ciencias Sociales y Humanidades CINTEOTL*, Universidad Autónoma del Estado de Hidalgo, No. 12, México.

Akerlof, G. A. (1970) “The market for ‘Lemons’: Quality uncertainty and the market mechanism” The Quarterly Journal of Economics, Vol. 84, No. 3, p 488-500.

Aleem, I. “Imperfect information, screening, and the costs of informal lending: A study of a rural credit market in Pakistan” *The World Bank Economic Review*, Vol. 4, No. 3 p 329-349.

Alkire, S. and Foster, J. (2011) “Understandings and misunderstandings of multidimensional poverty measurement” *OPHI Working Paper*, No. 43, Oxford Department of International Development, University of Oxford.

Alkire, S. Foster, J. Seth, S. Santos, M. E. Roche, J. M. and Ballon, P. (2015) *Multidimensional Poverty, Measurement and Analysis*, Oxford University Press, UK.

Allen, D. W. (1998) “No-fault divorce in Canada: Its cause and effect” *Journal of Economic Behaviour and Organization*, Vol. 37, p 129-149.

Altunbas, Y. Thornton, J. and Kara, A. (2010) “What determines financial exclusion? Evidence from Bolivian household data”, *Bangor Business School working paper*, BBSWP/10/018, UK.

Anderson, G. (2004) “Toward an empirical analysis of polarization” *Journal of Econometrics*, No. 122, p 1-26.

Angrist, J. D. and Pischke, J. S. (2009) *Mostly harmless econometrics*, Princeton University Press, US.

Aristotle (2017) *Politics*, Hackett Publishing Company, USA.

Armendariz, B. and Morduch, J. (2010) *The economics of microfinance*, second edition, Massachusetts Institute of Technology MIT, USA.

Arora, S. and Jain, K. (2006) "Testing for generalized Lorenz dominance" *Statistical Methods and Applications*, Vol. 15, p 75-88.

Ashenfelter, O. and Card, D. (1985) "Using the longitudinal structure of earnings to estimate the effects of training programs" *Review of Economics and Statistics*, Vol. 67, pp. 648-660.

Ashraf, N. (2009) "Spousal control and intra-household decision making: An experimental study in the Philippines", *American Economic Review*, Vol. 99, Issue 4, pp. 1245-1277.

Atkinson, A. B. (2003) "Multidimensional deprivation: Contrasting social welfare and counting approaches" *Journal of Economic Inequality*, Vol. 1, p 51-65.

Atkinson, A.B., L. Rainwater and T. M. Smeeding (1994) "Income distribution in European countries" *LIS Working Paper Series*, Working Paper No. 121, Luxembourg Income Study (LIS).

Atkinson, A. B. and Brandolini, A. (2011) "On the identification of the Middle Class", Working Paper Series, Society for the Study of Economic Inequality ECINEQ, WP 2011-217.

Atkinson, R. Parker, S. Burrows, R. (2016) "Elite formation, power and space in contemporary London" *Theory, Culture and Society*, Special issue: Elites and power after financialization, p 1-22.

Attanasio, O. and Szekely, M. (1999) "An asset-based approach to the analysis of poverty in Latin America" *Inter-American Development Bank*, working paper No. R-376.

Banerjee, A. V. and Duflo, E. (2008) "What is middle class about the middle classes around the world" *Journal of Economic Perspectives*, Vol. 22, No. 2, p 3-28.

Barcena, H. and Canto, O. (2018) "A simple subgroup decomposable measure of downward (and upward) income mobility" *ECINEQ*, Working Paper Series, 2018-472.

Barro, R. J. (1999) “Determinants of democracy” *Journal of Political Economy*, Vol. 107, No. 6, p 158-183.

Barros, R. (2013) “We were the middle class” *Latin America’s Middle Class*, Lexington Books, p 197-213.

Bateman, M. (2010) *Why doesn’t microfinance work?*, Zed Books, UK.

Beach, C. M. (1989) “Dollars and dreams: A reduced middle class? Alternative explanations” The Journal of Human Resources, Vol. 24, No. 1, p 162-193.

Becchetti, L, and Conzo, P. (2011) “Enhancing capabilities through credit access: Creditworthiness as a signal of trustworthiness under asymmetric information” Journal of Public Economics, Vol. 95, p 265-278.

Becker, G. (1976) *The economic approach to human behaviour*, Chicago University Press.

Beneria, L. and Roldan, M. (1987) *The Crossroads of Class and Gender*, The University of Chicago Press, p 31-50.

Bertrand, Marianne, Duflo, E. and Mullainathan, S. (2004) “How much should we trust differences-in-differences estimates?” *The Quarterly Journal of Economics*, Vol. 119, Issue 1, p 249-275.

Bester, H. (1985) “Screening vs rationing in credit markets with imperfect information” The American Economic Review, Vol. 75, No. 4 p 850-855.

Birdsall, N. (2010) “The (indispensable) middle class in developing countries; or, the rich and the rest, not the poor and the rest” *Center for Global Development*, working paper 207.

Blackburn, M. L. and Bloom, D. E. (1985) “What is happening to the middle class?” *American Demographics*, Vol. 7, No. 1, p 18-25.

Bloom, G. Standing, H. and Lloyd, R. (2008) “Markets, information asymmetry and health care: Thowards new social contracts” *Social Science and Medicine*, Vol. 66, p 2067-2087.

Blundell, R. and Costa-Dias, M. (2009) “Alternative approaches to evaluation in empirical microeconomics” *The Journal of Human Resources*, Vol. 44, No. 3, p 565-640.

Boyle, D. (2013) *Broke, who killed the middle class*, Fourth Estate, London.

Bourdieu, P. (1986) “The Forms of Capital”, *Handbook of Theory and Research for the Sociology of Education*, John G. Richardson, ed. Westport, CT: Greenwood Press, p 241-260.

Bourguignon, F. and Chiappori, P (1992) “Collective models of household behaviour, an introduction” *European Economic Review*, No. 36, p 355-364.

Bradley, R. A. and Terry, M. E. (1952) “Rank analysis of incomplete block designs: I. The method of paired comparisons” *Biometrika*, Vol. 39, No. ¾, p. 324-345.

Caliendo, M. and Kopeinig, S. (2008) “Some practical guidance for the implementation of propensity score matching” *Journal of Economic Surveys*, Vol. 22, No. 1, p 31-72.

Card, D. and Krueger, A. B. (1994) “Minimum wages and employment: A case study of the fast-food industry in New Jersey and Pennsylvania” *The American Economic Review*, Vol. 84, No. 4, p 772-793.

Castellani, F. Parent, G. and Zenteno-Gonzales, J. (2015) “Who is the Latin American middle class? Relative-income and multidimensional approaches” *Latin America’s emerging middle classes economic perspectives*, International Political Economy Series, Palgrave Macmillan, p 101-129.

Cazenave, N. A. (1983) “A woman’s place: the attitudes of the middle class black men” *Phylon*, Vol. 44, No. 1, p12-32.

Chakravarty, S. R. (2015) *Inequality, bipolarization and conflict, an analytical study*, Springer India, Economic Studies in Inequality, Social Exclusion and Well-being series, p 1-50.

Chant, S (1991) *Women and Survival in Mexican Cities*, Manchester University Press, p 143-170.

Cherlin, A. Cross-Barnet, C. Burton, L. M. Garrett-Peters, R. (2008) “Promises they can keep: low-income women’s attitudes toward motherhood, marriage, and divorce” *Journal of Marriage and Family*, No. 70, pp 919-933.

Chong, A. and La Ferrara, E. (2009) “Television and divorce: evidence from Brazilian *novelas*” *Journal of the European Economic Association*, No. 7, Vol. 2-3, p 458-468.

CIF (2009) *Perspectivas para México de inclusión financiera integral*, Informe oficial del proyecto de inclusión financiera en 2020 del Centro para la Inclusión Financiera en ACCIÓN internacional, Centro para la Inclusión Financiera, México.

CNDH (2013) *Cuarta Visitaduria General, Programa de Asuntos de la Mujer y de Igualdad entre Mujeres y Hombres*, mimeo, Comisión Nacional de Derechos Humanos, México.

CNIF (2012) 4º *Reporte de Inclusión Financiera*, Consejo Nacional de Inclusión Financiera, México.

CNIF (2016) *Política Nacional de Inclusión Financiera*, Consejo Nacional de Inclusión Financiera, México.

Cole, G. D. H. (1950) "The conception of the middle classes" *The British Journal of Sociology*, Vol. 1, No. 4, p 275-290.

Coleman, J.S. (1988) "Social Capital in the Creation of Human Capital", *American Journal of Sociology*, Vol 94, Supplement, p S95-S120.

Colletta, N. J. and Cullen, M. L. (2000) *"The nexus between violent conflict, social capital and social cohesion: case studies from Cambodia and Rwanda"* Woking Paper No. 23, The World Bank.

CONASAMI (2017) *Salarios mínimos generales por área geográfica*, mimeo, Comisión Nacional de los Salarios Mínimos, México.

CONEVAL (2012) *Construcción de las Líneas de Bienestar*, Documento metodológico, Consejo Nacional para la Evaluación de la Política de Desarrollo Social CONEVAL, primera edición, México.

CONEVAL (2013) *Apoyos para la Inclusión Financiera y la Bancarización*, Ficha de Monitoreo 2013, Consejo Nacional de Evaluación de la Política de Desarrollo Social, México.

CONEVAL (2018) *Metodología para la medición multidimensional de la pobreza en México*, Consejo Nacional para la Evaluación de la Política de Desarrollo Social CONEVAL, tercera edición, México.

Cormier, D. Aerts, W. Ledoux, M. J. and Magnan, M. (2009) "Attributes of social and human capital disclosure and information asymmetry between managers and investors" *Canadian Journal of Administrative Sciences*, Vol. 26, p 71-88.

Correa-Cabrera, G. (2012) "The spectacle of drug violence: American public discourse, media and border enforcement in the Texas-Tamaulipas border region during drug-war times" *Norteamerica, Revista Academica del CSIAN-UNAM*, Vol. 7, No. 2, p 199-220.

Cowell, F. A. (2013) *Measuring Inequality*, third edition, Oxford University Press, London, p 18-37.

Crompton, R. (2008) *Class and Stratification*, Polity, 3rd edition, UK.

Currid-Halkett, E. (2017) *The sum of small things, a theory of the aspirational class*, Princeton University Press, USA.

Cutrona, C. E. and Russell, D. W. (1987) “The provisions of social relationships and adaptation to stress” *Advances in Personal Relationships*, Vol. 1, p 37-67.

DANE (2018) *Pobreza Monetaria y Multidimensional en Colombia 2017*, Technical bulletin, National Administrative Department of Statistics, Bogota, Colombia.

Daniel-Rosen, J. and Zepeda-Martinez, R. (2015) “La guerra contra el narcotrafico en Mexico: Una Guerra perdida” *Reflexiones*, Vol. 94, No. 1, p 153-168.

Dang, H. H. and Lanjouw, P. F. (2013) “Measuring poverty dynamics with synthetic panels based on cross-sections” *Policy Research Working Paper Series*, No. 6504, The World Bank, Washington DC.

Davis, H. L. (1976) “Decision making within the household” *Journal of Consumer Research*, Vol. 2, No. 4, p 241-260.

Davis, J. C. and Huston, J. H. (1992) “The shrinking middle-income class: A multivariate analysis” *Eastern Economic Journal*, Vol. 18, No. 3, p 277-285.

Dayton-Johnson, J. (2015) “Making sense of Latin America’s middle classes” *Latin America’s emerging middle classes economic perspectives*, International Political Economy Series, Palgrave Macmillan, p 1-31.

Dardanoni, V. and Forcina, A. (1999) “Inference for Lorenz curve orderings” *Econometrics Journal*, Vol. 2, p 49-75.

Daza, N. A. and Cortes, M. F. (2013) “Measurement and characterization of the middle class in Latin America” *Archivos de Economia*, Direccion de Estudios Economicos, documento 405, Colombia.

De la Calle, L and Rubio, L. (2010) *Clasemediero*, Centro de Investigación para el Desarrollo A. C., México.

Deaton, A. (1985) “Panel data from time series of cross-sections” *Journal of Econometrics*, Vol. 30, p 109-126.

Deaton, A. and Muellbauer, J. (1999) *Economics and consumer behaviour*, Cambridge University Press, 13a reprint, USA.

Dehejia, R. and Wahba, S. (2002) “Propensity score-matching methods for nonexperimental causal studies” *The Review of Economics and Statistics*, Vol. 84, No. 1, p 151-161.

Deree, C. D. Alvarado, G. E. Twyman, J. (2012) “Gender inequality in asset ownership in Latin America: Female owners vs household heads”

Development and Change, Vol. 43, No. 2, p 505-530. Devlin, J. F. (2005) “A detailed study of financial exclusion in the UK” *Journal of Consumer Policy*, Vol. 28, p 75-108.

Doepek, M. and Zilibotti, F. (2005) “Social class and the spirit of capitalism” *Journal of the European Economic Association*, Vol. 3, No. 2/3, p 516-524.

Doll, C. N. H. Muller, J-P and Elvidge, C. D. (2000) “Night-Time Imagery as a Tool for Global Mapping of Socioeconomic Parameters and Greenhouse Gas Emissions” *Ambio*, 29 (3) (2000), p 157-162.

Durlauf, S. N. and Fafchamps, M. (2004) “Social Capital”, *NBER working papers series*, National Bureau of Economic Research, working paper 10485.

Dymski, G. (2014) “Analytical aspects of financial access and financial exclusion” *Development on Trial*, Orient Black Swan, India.

Easterly, W. (2001) “The middle class consensus and economic development” Journal of Economic Growth, Vol. 6, p 317-335.

EC (2008) *Financial services provision and prevention of financial exclusion*, Directorate-General for employment, social affairs and equal opportunities, The European Commission.

Edo, M. Sosa-Escudero, W. and Svarc, M. (2017) *“Multidimensional middle class”* Seventh meeting of the Society for the Study of Economic Inequality, ECINEQ, July 17-19 2017, paper 86.

Eisenhauer, J. (2008) “An economic definition of the middle class” *Forum of Social Economics*, Vol. 37, No. 2, p 103-113.

ENIF (2015) *Encuesta Nacional de Inclusión Financiera 2015*, Libro de Tabulados, INEGI y la Comisión Nacional Bancaria y de Valores (CNBV), México.

Estache, A. and Leipziger, D. (2009) *Stuck in the middle: Is fiscal policy failing the middle class?* Brookings Institution Press, USA.

Esteban, J-M and Ray, D. “On the measurement of polarization” *Econometrica*, Vol. 62, No.4, p 819-851.

Etter, G. W. and Lehmuth, E. L. (2013) “The Mexican drug wars: Organized crime, narco-terrorism, insurgency or asymmetric warfare?” *Journal of Gang Research*, Vol. 20, No. 4, p 1-34.

FCL (2013) *Codigo Civil Federal*, Federal Civil Law, last reform in 2013, Mexico.

Ferreira, F. H. G. Messina, J. Rigolini, J. Lopez-Calva, L. F. Lugo, M. A. Vakis, R. (2013) *Economic Mobility and the Rise of the Latin American Middle Class*, The World Bank, Washington DC.

Filmer, D. and Pritchett, L. H. “Estimating wealth without expenditure data-or tears: an application to educational enrolment in States of India” *Demography*, Vol. 38, No. 1, p 115-132.

Floro, M. S. and Ray, D. (1997) “Vertical links between formal and informal financial institutions” *Review of Development Economics*, Vol. 1, No. 1, p 34-56.

Forster, M. (1994) “Measurement of low incomes and poverty in a perspective of international comparisons” *Labour Market and Social Policy Occasional Papers*, No. 14, Organisation for Economic Co-Operation and Development, OECD.

Foster, J. Seth, S. Lokshin, M. and Sajaia, Z. (2013) *Measuring poverty and inequality: theory and practice*, The World Bank, USA.

Foster, J. and Wolfson, M. C. (1992) “Polarization and the decline of the middle class: Canada and the U.S.” *OPHI Working Paper*, No. 31, Oxford Department of International Development, University of Oxford.

Foster, J. and Wolfson, M. C. (2010) “Polarization and the decline of the middle class: Canada and the U.S.” *The Journal of Economic Inequality*, Vol. 8, Issue 2, p 247-273.

French, W. E. (2013) “Moralizing the masses” *Latin America’s Middle Class*, Lexington Books, p 79-104.

Frey B. S. (1997) *Not Just for the Money*, Edward Elgar Publishing Limited, UK.

Friedman, M (1957) *A Theory of the Consumption Function*, Princeton University Press, USA.

Friedrich, P. (1965) “A Mexican Cacicazgo” *Ethnology*, Vol. 4, No. 2, p 190-209.

Gallegos, J. V. and Ondrich, J. I. (2017) “The effects of the Chilean divorce law on women’s first birth decisions” *Review of Economics of the Household*, Vol. 15, Issue 3, p 857-877.

Gehrig, T. (1998) “Screening, cross-border banking, and the allocation of credit” *Research in Economics*, Vol. 52, p 387-407.

Genadek, K. R. Stock, W. A. and Stoddard, C. (2007) “No-fault divorce laws and the labour supply of women with and without children” *Journal of Human Resources*, Vol. 42, Issue 1, p 247-274.

Gertler, P. Martinez, S. Premand, P. Rawlings, L. Vermeersch, C. (2011) *Impact Evaluation in Practice*, The World Bank, Washington DC.

Glennerster, H. Lupton, R. Noden, P. and Power, A. (1999) “Poverty, social exclusion and neighbourhood: Studying the area bases of social exclusion” *LSE STICERD*, research paper No. CASE 022, London School of Economics, UK.

Gordon, D. Levitas, R. Pantazis, C. Patsios, D. Payne, S. and Townsend, P. (2000) *Poverty and Social Exclusion in Britain*, Joseph Rowntree Foundation, UK.

Graham, C. and Pettinato, S. (2001) “Happiness, markets, and democracy: Latin America in comparative perspective” *Journal of Happiness Studies*, Vol. 2, p 237-268.

Granovetter, M. S. (1973) “The strength of weak ties” *American Journal of Sociology*, Vol. 78, No. 6, p 1360-1380.

Granovetter, M. (1985) “Economic action and social structure: The problems of embeddedness” *American Journal of Psychology*, Vol. 91, No. 3, p 481-510.

Greene, W. H. (2012) *Econometric Analysis*, seventh edition, Pearson.

Grootaert, C. Narayan, D. Jones, V. N. Woolcock, M. (2004) “Measuring Social Capital” *World Bank Working Paper*, No. 18, The World Bank.

Grossbard, S. (2010) “Independent individual decision-makers in household models and the new home economics” *IZA Discussion Paper*, No. 5138.

Gutin, G. Yeo, A. Zverovich, A. (2002) “Traveling salesman should not be greedy: domination analysis of greedy-type heuristics for the TSP” *Discrete Applied Mathematics*, Vol. 117, p 81-86.

Heckman, J. J. Ichimura, H. and Todd, P. E. (1997) “Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme” *Review of Economic Studies*, Vol. 64, p 605-654.

Heckman, J. J. Ichimura, H. and Todd, P. E. (1998) “Matching as an econometric evaluation estimator” *Review of Economic Studies*, Vol. 65, p 261-294.

Hill-Kay, H. (1987) “Equality and difference: A perspective on no-fault divorce and its aftermath” *University of Cincinnati Law Review*, Vol. 56, No. 1, p 1-90.

Hochschild, A. and Machung, A. (2012) *The Second Shift, Working parents and the revolution at home*, Penguin Publishing Group, London.

Hoff, K. and Stiglitz, J. E. “Imperfect information and rural credit markets: Puzzles and policy perspectives” *The Economics of Rural Organizations*, The World Bank.

Holvoet, N. (2005) “The impact of microfinance on decision-making Agency: Evidence from south India” *Development and Change*, Vol. 36, Issue 1, p 75-102.

Horney, M. J. and McElroy, M. B. (1988) “The household allocation problem: empirical results from a bargaining model” *Research in Population Economics*, Vol. 6, p 15-38.

House, J. S. Landis, K. R and Umberson, D. (1988) “Social relationships and health” *Science*, Vol. 241, No. 4865, p 540-545.

ILO (2014) “Informal employment in Mexico: current situation, policies and challenges” *Notes on Formalisation*, Programme for the Promotion of Formalisation in Latin America and the Caribbean, International Labour Organisation, Regional Office for Latin America and the Caribbean.

INEGI (2000) *Uso del tiempo y aportaciones en los hogares mexicanos*, Instituto Nacional de Estadística y Geografía, INEGI, mimeo, México.

Jacob, H. (1989) “Another look at no-fault divorce and the post-divorce finances for women” *Law and Society Review*, Vol. 23, No. 1, p 95-116.

Johnson, J. H. and Mazingo, C. J. (2000) “The economic consequences of unilateral divorce for children” *University of Illinois CBA Office of Research Working Paper*, No. 00-0112.

Johnson, W. R. and Skinner, J. (1986) “Labor supply and marital separation” *The American Economic Review*, Vol. 76, No. 3, p 455-469.

Kempson, E. and Whyley, C. (1999) *Keep out or opted out? Understanding and combating financial exclusion*, Bristol: Policy press, UK.

Kenny, C. (2011) “Where is the virtue in the middle class?” *Center for Global Development*, essay 1425646.

Kneip, T. and Bauer, G. (2009) “Did unilateral divorce laws raise divorce rates in Western Europe?” *Journal of Marriage and Family*, Vol. 71, Issue 3, p 592-607.

LACP (2001) *Ley de Ahorro y Crédito Popular*, Cámara de Diputados del H. Congreso de la Unión, México.

Lapavitsas, C. (2007) “Information and trust as social aspects of credit” *Economy and Society*, Vol. 36, No. 3, p 416-436.

Layte, R. Maitre, B. Nolan, B. Watson, D. Williams, J. and Casey, B. (2000) “Monitoring poverty trends: Results from the 1998 living in Ireland survey” *ESRI working paper*, Economic and Social Research Institute, No. 132.

Levy, F. (1987) “The middle class: is it really vanishing?” The Brookings Review, Vol. 5, No. 3, p 17-21.

Leyshon, A. and Thrift, N. (1995) “Geographies of financial exclusion: financial abandonment in Britain and the United States”, *Transactions of the Institute of British Geographers, New series*, 20, p 312-341.

Lin, M. Prabhala, N. R. and Viswanathan, S. (2013) “Judging borrowers by the company they keep: Friendship networks and information asymmetry in online peer-to-peer lending” *Management Science*, Vol. 54, No. 1, p 17-35.

List, J. A. Millimet, D. L. Fredriksson, P. G. and McHone, W. W. (2003) “Effects of environmental regulations on manufacturing plant births” *The Review of Economics and Statistics*, Vol. 85, No. 4, p 944-952.

Loayza, N. Rigolini, J. and Llorente, G. (2012) “Do middle classes bring about institutional reforms?” *Economic Letters*, Vol. 116, p 440-444.

Lopez, A. R. (2013) “It is not something you can be or come to be overnight” *Latin America’s Middle Class*, Lexington Books, p 151-170.

Lopez-Calva, L. F. and Ortiz-Juarez, E. (2014) “A vulnerability approach to the definition of the middle class” Journal of Economic Inequality, Vol. 12, p 23-47.

Lopez-Rodriguez, P. and De la Torre, R. (2010) “Social Capital in the Presence of Market Failures”, *International Journal of Social Inquiry*, Vol. 3, No. 1, p 163-188.

Luke, N. and Munshi, K. (2006) “New roles for marriage in urban Africa: Kinship networks and the labor market in Kenya” *Review of Economics and Statistics*, Vol. 88, No. 2, p 264-282.

Lyons, A. C. (2003) “How credit access has changed over time for US households”, *The Journal of Consumer Affairs*, Vol. 37, No. 2, p 231-255.

Maasoumi, E. and Yalonenetzky, G. (2013) “Introduction to robustness in multidimensional wellbeing analysis” *Econometric Reviews*, Vol. 32, No. 1, p 1-6.

Mansell-Carstens, C. (1995) *Las finanzas populares en México*, ITAM, México.

Mehta, J. S. and Gurland, J. (1969) “Testing equality of means in the presence of correlation” *Biometrika*, Vol. 56, No. 1, p 119-126.

Manser, M. and Brown, M. (1980) “Marriage and household decision-making: a bargaining analysis”, *International Economic Review*, Vol. 21, No. 1, p 31-44.

Manski, C. F. (1995) *Identification problems in the social sciences*, Harvard College, US.

Marco-Navarro, F. (2009) “Legislación comparada en material de familias. Los casos de cinco países de América Latina” *Serie Políticas Sociales,* No. 149, CEPAL.

Marx, K. (2015) *Capital*, Progress Publishers, Vol. 1, Proofed version 2015.

Mass-Collel, A. Whinston, M. D. and Green, J. R. (1995) *Microeconomic Theory*, Oxford University Press, USA.

Meyer, B. D. and Sullivan, J. X. (2011) “Consumption and income poverty over the business cycle” *NBER Working Paper*, No. 16751, NBER, USA.

Milanovic, B. and Yitzhaki, S. (2002) “Decomposing world income distribution: does the world have a middle class?” Review of Income and Wealth, Series 48, No. 2, p 155-178.

Miller, A. J. and Sassier, S. (2011) “The spectre of divorce: views from working –and middle class-cohabiters” *Family Relations*, No. 60, p 602-616.

Morales-Oyarvide, C. (2011) “La Guerra contra el narcotráfico en México. Debilidad del Estado, orden local y fracaso de una estrategia” *Aposta Revista de Ciencias Sociales*, No. 50, p 1-35.

Morrison, D. F. (1973) “A test for equality of means of correlated variates with missing data on one response” *Biometrika*, Vol. 60, No. 1, p 101-105.

Moser, C. O. N. and Holland, J. (1997) *Urban poverty and violence in Jamaica*, World Bank Latin American and Caribbean Studies, The World Bank, US, p 5-22.

Mwangi, I. W. and Ouma, S. A. (2012) “Social capital and access to credit in India” American Journal of Social and Management Sciences, Vol. 3 No. 1, p 8-16.

Nakonezny, P. A Shull, R. D. and Rodgers, J. L. (1995) “The effect of no-fault divorce law on the divorce rate across the 50 States and its relation to income, education and religiosity” *Journal of Marriage and Family*, Vol. 57, No. 2, p 477-488.

Nussbaum, M. C. (2011) *Creating capabilities, the human development approach*, Harvard University Press, USA.

OECD (2010) *Latin America economic outlook 2011: How middle-class is Latin America?*, Organization for Economic Co-operation and Development.

OECD (2019) *What are equivalence scales?* OECD Project on Income Distribution and Poverty, Organisation for Economic Co-Operation and Development, mimeo.

Ostrom, E. (1999) “Social capital: a fad or a fundamental concept?” *Social Capital, A Multifaceted Perspective*, The World Bank, p 172-213.

Ostrom, E. (2000) “Collective action and the evolution of social norms” *Journal of Economic Perspectives*, Vol. 14, No. 3, p 137-158.

Owensby, B. P. (2013) “Domesticating modernity: Markets, home, and morality in the middle class in Rio de Janeiro and Sao Paulo, 1930s and 1940s” *Latin America’s Middle Class*, Lexington Books, p 129-150.

Parker, D. S. (2013) “The making and endless remaking of the middle class” *Latin America’s Middle Class*, Lexington Books, p 1-22.

Parkman, A. M. (1992) *No-fault Divorce: What Went Wrong?* Westview Press, US.

Partridge, M. D. (1997) “Is inequality harmful for growth? Comment” *The American Economic Review*, Vol. 87, No. 5, p 1019-1032.

Pedrero, M. (2004) “Genero, trabajo doméstico y extra-doméstico en México. Una estimación del valor económico del trabajo doméstico” *Estudios Demográficos y Urbanos*, Vol. 19, No. 2 (56), p 413-446.

Peters, H. E. (1986) “Marriage and divorce: informational constraints and private contracting” *The American Economic Review*, Vol. 76, No. 3, p 437-454.

Pike, F. B. (2013) “Aspects of class relations in Chile, 1850-1960” *Latin America’s Middle Class*, Lexington Books, p 25-34.

Piketty, T. (2014) *Capital in the twenty-first century*, Harvard University Press.

Piketty, T. and Saez, E. (2003) “Income inequality in the United States, 1913-1998” The Quarterly Journal of Economics, Vol. 158, Issue 1, p 1-39.

Putnam, R. D. (2000) *Bowling Alone The collapse and revival of American Community*, Simon and Schuster paperbacks, USA.

Quah, D. T. (1996) “Twin Peaks: Growth and Convergence in Models of Distribution Dynamics” *The Economic Journal*, Royal Economic Society, Vol. 106, Issue 437, p 1045-1055.

Ravallion, M. (2010) “The developing world’s bulging (but vulnerable) middle class” *World Development*, Vol. 38, No. 4, p 445-454.

Ravallion, M., Dutta, P., Murgai, R., and van de Walle, D. (2014) *Right to work? Assessing India´s Employment Guarantee Scheme in Bihar,* The World Bank, Washington DC.

Ray, D. (1998) *Development economics*, Princeton University Press, USA.

Robeyns, I. (2005) “The capability approach: A theoretical survey” *Journal of Human Development*, Vol. 6, No. 1, p 93-117.

Reves, R. V. Guyot, K. and Krause, E. (2018) “Defining the middle class: Cash, credentials, or culture?” *The Future of the Middle Class Initiative*, Report, The Brookings Institution, May, 2018.

Robison, L. J. and Siles, M. E. (2012) “Introducción al capital social y su paradigma”, *Capital social y política pública en México*, Patricia López-Rodríguez and Isidro Soloaga (editores), El Colegio de México, México.

Rodriguez, J. G. and Salas, R. (2003) “Extended bi-polarization and inequality measures” *Research on Economic Inequality*, Vol. 9, p 69-83.

Rodriguez-Oreggia, E. Lima, M. and Mulato-Martínez, D. (2013) “Competencia de Mercado e informalidad en el sector manufacturero mexicano, 1987-2004” *Economía, Sociedad y Territorio*, Vol. 13, No. 43, p 721-750.

Romer, D. (2012) *Advanced Macroeconomics*, McGraw-Hill Irwin, fourth edition.

Rosenbaum, P. R. and Rubin, D. B. (1983) “The central role of the propensity score in observational studies for causal effects” *Biometrika*, Vol. 70, No. 1, p 41-55.

Saarimaki, P. (2018) “Bourgeois women and the question of divorce in Finland in the late 19th and 20th centuries” Scandinavian Journal of History, Vol. 43, No. 1, p 64-90.

Sandoval, H. and Lima, M. (2014) “El capital social en el acceso al mercado laboral y la pobreza en México”, *Bienestar y política social*, vol. 9, no. 1, p 53-78.

Savage, M. (2015) *Social class in the 21st century*, Pelican Books, UK.

SCJN (2012) “Trámite procesal del divorcio sin expresión de causa” *Cuadernos de Trabajo de la Primera Sala*, Suprema Corte de Justicia de la Nación, No. 4, México.

Sebastian, A. (2011) “Is divorce a credible threat? Unilateral separation laws, household bargaining and women´s labour in Egypt”, *mimeo*, University of Maryland, College Park.

Seibel, H. D. (2001) “Mainstreaming informal financial institutions” *Working Paper*, No. 2001.6, University of Cologne Development Research Centre, Germany.

SEP (2018) *La Estructura del Sistema Educativo Mexicano*, Dirección General de Acreditación, Incorporación y Revalidación, Secretaria de Educación Pública SEP, México.

Serrano, M. (2010) “El problema del narcotráfico en México: Una perspectiva latinoamericana” *Los Grandes Problemas de México XIII, Relaciones Internacionales*, El Colegio de México, México.

SHCP (2013) *Programa Nacional de Financiamiento del Desarrollo*, Secretaria de Hacienda y Crédito Público, México.

Sianesi, B. (2014) “Dealing with randomisation bias in a social experiment: The case of ERA”, *Institute for Fiscal Studies*, IFS working paper W14/10.

Sinclair, S. (2001) “Financial exclusion: an introductory survey” *Centre for Research into Socially Inclusive Services*, Edinburgh.

Smith, J. P. and Ward, M. P. (1985) “Time-series growth in the female labour force” *Journal of Labor Economics*, Vol. 3, No. 1, p S59-S90.

Solimano, A. (2008) “The middle class and the development process” *Serie Macroeconomia del Desarrollo*, No. 65, United Nations and Economic Commission for Latin America and the Caribbean, Chile.

Stiglitz, J. and Weiss, A. (1981) “Credit rationing in markets with imperfect information” *The American Economic Review*, Vol.71, No. 3, p 393-410.

Stoyanova, S. and Tonkin, R. (2018) “An expenditure-based approach to poverty in the UK” 35th *IARIW General Conference*, Copenhagen Denmark.

Temin, P. and Vines, D. (2013) *The Leaderless Economy*, Princeton University Press, US, p 1-20.

Thomas, D. (1990) “Intra-household resource allocation: An inferential approach” *The Journal of Human Resources*, Vol. 25, No. 4, p 635-664.

Thurow, L. C. (1984) “The disappearance of the middle class”, *New York Times*, F.3.

Townsend, P. (1979) *Poverty in the United Kingdom*, Penguin Books, UK.

Valenzuela, A. (1999) “Gender roles and settlement activities among children and their immigrant families” *American Behavioural Scientist*, Vol. 2, No. 4, p 720-742.

Varoufakis, Y. (2013) *The Global Minotaur*, second edition, Zed Books, London.

Veblen, T. (2009) *The theory of the leisure class*, Oxford University Press, first reissued.

Villas-Boas, J. M. and Schmidt-Mohr, U. (1999) “Oligopoly with asymmetric information: Differentiation in credit markets” *The RAND Journal of Economics*, Vol. 30, No. 3, p 375-396.

Vranken, J. (2002) “Belgian reports on poverty” *Reporting on Income Distribution and Poverty – Perspectives from a German and European point of view*, conference paper, Berlin.

Wagley, C. (2013) “The dilemma of the Latin America middle class” *Latin America’s Middle Class*, Lexington Books, p 65-77.

Wang, Y. and Tsui, K. (2000) “Polarization Orderings and New Classes of Polarization Indices”, *Journal of Public Economic Theory*, Vol. 2, No. 3, p 349-363.

Weber, M. (2001) *The Protestant Ethic and the Spirit of Capitalism*, Routledge Classics, reprinted, USA.

Wentzel, K. R. (1998) “Social relationships and motivation in middle school: the role of parents, teachers and peers” *Journal of Educational Psychology*, Vol. 90, No. 2, p 202-209.

Wilks, S. S. (1946) “Sample criteria for testing equality of means, equality of variances, and equality of covariance in a normal multivariate distribution” *The Annals of Mathematical Statistics*, Vol. 17, No. 3, p 257-281.

Wheatley, D. and Wu, Z. (2014) “Dual careers, time-use and satisfaction levels: evidence from the British Household Panel Survey” *Industrial Relations Journal*, Vol. 45, Issue 5, p 443-464.

Whiteford, A. H. (2013) “Community pillars: The middle class” *Latin America’s Middle Class*, Lexington Books, p 45-54.

Williams, L. and Guest, M. P. (2004) “Attitudes toward marriage among the urban middle class in Vietnam, Thailand and the Philippines” *Journal of Comparative Family Studies*, Vol. 36, No. 2, p 163-186.

Woolcock, M. (1999) “Managing risk, shocks and opportunity in developing economies: the role of social capital”, *Dimensions of development*, G. Ranis (editor), New Haven: Yale Center for International and Area Studies.

Wooldridge, J. M. (2002) *Econometric Analysis of Cross Section and Panel Data*, second edition, MIT press.

Wu, W-P (2008) “Dimensions of social capital and firm competitiveness improvement: The mediating role of information sharing” *Journal of Management Studies*, Vol. 41, No. 1, p 122-146.

Yalonetzky, G. (2014) “Relative Bipolarization Lorenz Curve”, *Economics Bulletin*, Vol. 34, Issue 2, p 1103-1113.

Yalonetzky, Gaston (2016) “Coherent measurement of relative bipolarisation with normalised median-independent indices”, mimeograph.

Yalonetzky, Gaston (2017) “The benchmark of maximum relative bipolarisation” *Research on Economic Inequality*, Vol. 25, p 51-62.

Zeller, M. Schrieder, G. von Braun, J. Heidhues, F. (1997) *Rural Finance for Food Secrity for the Poor, Implications for Research and Policy*, International food policy research institute, Washington, DC., p 6-12.

List of Abbreviations

AL	Affluent Line
ATE	Average Treatment Effect
ATM	Automatic Teller Machine
ATNT	Average Treatment effect on the Non-Treated
ATT	Average Treatment effect on the Treated
CASEN	Encuesta de Caracterización Socioeconómica Nacional
CEO	Chief Executive Officer
CNDH	Comisión Nacional de los Derechos Humanos
CNIF	Consejo Nacional para la Inclusión Financiera
CONEVAL	Consejo Nacional de Evaluación de la Política de Desarrollo Social
DANE	Departamento Administrativo Nacional de Estadística
DF	Distrito Federal
DGEEC	Dirección General de Estadística, Encuestas y Censos
DID	Difference-in-Difference
DVD	Digital Versatile Disc
EC	European Commission
ECH	Encuesta Continua de Hogares
ENAHO	Encuesta Nacional de Hogares sobre Condiciones de vida y Pobreza
ENCAS	Encuesta Nacional de Capital Social
ENIF	Encuesta Nacional de Inclusión Financiera
ENIGH	Encuesta Nacional de Ingresos y Gastos de los hogares
ENOE	Encuesta Nacional de Ocupación y Empleo
EPH	Encuesta Permanente de Hogares
EPL	Extreme Poverty Line
FCL	Federal Civil Law
FW	Foster and Wolfson
GDP	Gross Domestic Product
GP	General Practitioner
HEI	Higher-Education Index
IBGE	Instituto Brazileiro de Geografia e Estadistica
IMSS	Mexican Institute for Social Security
INE	Instituto Nacional de Estadística
INEGI	Instituto Nacional de Estadística y Geografía
INEI	Instituto Nacional de Estadística e Informática
ISSSTE	Mexican Institute for Social Security and Services for State Workers
MCS-ENIGH	Módulo de Condiciones Socioeconómicas de la Encuesta Nacional de Ingresos y Gastos de los hogares

MPI	Multidimensional Poverty Index
MSD	Ministerio de Desarrollo Social
OECD	Organisation for Economic Co-operation and Development
PEMEX	Petróleos Mexicanos
PGR	Procuraduría General de la Republica
PL	Poverty Line
PNAD	Pesquisa Nacional Por Amostra de Domicilios
PPP	Purchase Power Parity
PSM	Propensity Score Matching
RBL	Relative Bipolarisation Lorenz
RCS	Repeated Cross-Section
SCJN	Suprema Corte de Justicia de la Nación
SE	Standard Error
SEDENA	Secretaria de la Defensa
SEMAR	Secretaria de Marina
SEP	Mexican Ministry of Education
SHCP	Secretaria de Hacienda y Crédito Publico
SP	Seguro Popular
SS	Social Security
TV	Television
UNAM	Universidad Nacional Autónoma de México
UNDP	United Nations Development Program
USA	United States of America
USD	United States Dollar
VCR	Videocassette Recorder
WT	Wang and Tsui

Appendix A

A.1 Real monthly per capita income by country, descriptive statistics

A.1.1 Real* monthly per capita income, Brazil 2004 – 2015

Year	Observations	Mean	Median	Minimum	Maximum	Gini
2004	110351	639.19	349.32	0.00	82291.39	0.5781
2005	114576	664.16	370.87	0.00	81129.44	0.5747
2006	116333	708.60	398.21	0.00	95811.06	0.5667
2007	115112	731.65	424.97	0.00	69857.60	0.5595
2008	114821	780.17	457.22	0.00	93647.10	0.5535
2009	117827	799.43	474.17	0.00	99439.81	0.5495
2011	106447	842.73	511.08	0.00	65643.65	0.5346
2012	110501	901.75	553.39	0.00	73933.36	0.5272
2013	111029	929.63	567.97	0.00	68692.77	0.5270
2014	115019	962.78	570.41	0.00	72482.41	0.5212
2015	115624	901.81	569.41	0.00	51786.52	0.5197

Source: own estimations using microdata from PNAD, 2004-2015.
*Real income estimation is based on Consumer Price Index (CPI) per year, published by the World Bank (World Bank, 2018). Base year 2010.

A.1.2 Real* monthly per capita income, Chile 2000 – 2011

Year	Observations	Mean	Median	Minimum	Maximum	Gini
2000	64650	96836.65	47454.88	56.67	49184463.26	0.6086
2003	68153	162577.20	90681.78	0.00	50681465.96	0.5451
2006	73720	181916.50	107230.94	0.00	37113801.23	0.5192
2009	71460	187033.68	119477.15	0.00	15358742.16	0.4824
2011	86854	249144.39	146620.22	0.00	44446134.04	0.5133

Source: own estimations using microdata from CASEN, 2000-2011.
*Real income estimation is based on Consumer Price Index (CPI) per year, published by the World Bank (World Bank, 2018). Base year 2010.

A.1.3 Real* monthly per capita income, Mexico 2000 – 2014

Year	Observations	Mean	Median	Minimum	Maximum	Gini
2000	10108	3202.17	1948.33	62.14	155469.49	0.5197
2002	17167	3213.26	1988.87	69.93	1055396.96	0.5073
2004	22595	4246.23	2425.98	51.35	1538116.11	0.5420
2005	23174	3864.31	2254.29	0.00	414515.86	0.5331
2006	20875	4121.77	2465.38	0.00	211667.60	0.5188
2008	29468	4280.79	2517.18	0.00	1590362.27	0.5269
2010	27655	3603.88	2243.84	0.00	221804.33	0.5025
2012	9002	3591.21	2172.98	70.98	245495.67	0.5039
2014	19479	3544.19	2223.23	0.00	297523.23	0.4865

Source: own estimations using microdata from ENIGH, 2000-2014.
*Real income estimation is based on Consumer Price Index (CPI) per year, published by the World Bank (World Bank, 2018). Base year 2010.

A.1.4 Real* monthly per capita income, Paraguay 2002 – 2015

Year	Observations	Mean	Median	Minimum	Maximum	Gini
2002	3789	883114.56	476182.30	12849.73	124981208.11	0.5805
2003	9591	885848.47	462216.96	0.00	213934171.09	0.5830
2004	7823	873449.68	485699.99	0.00	109447779.21	0.5545
2005	4464	960286.96	558219.65	0.00	35703381.12	0.5316
2006	5292	914318.57	513041.17	0.00	277626463.37	0.5541
2007	4812	1053475.41	563371.83	0.00	300040731.69	0.5773
2008	4601	993316.61	564017.62	5846.13	74210813.25	0.5417
2009	4439	962934.17	579069.42	0.00	28002796.74	0.5270
2010	5003	1147893.00	657147.00	16221.00	183500000.00	0.5560
2011	4894	1131407.95	685925.99	0.00	98472140.76	0.5315
2012	5288	1146360.34	734912.66	29504.84	55879267.56	0.5017
2013	5424	1279231.10	811921.75	0.00	65808319.16	0.5035
2014	5165	1300590.98	802043.28	20503.80	143835590.12	0.5122
2015	8229	1199468.71	780804.36	0.00	84355837.47	0.4910

Source: own estimations using microdata from EPH, 2002-2015.
*Real income estimation is based on Consumer Price Index (CPI) per year, published by the World Bank (World Bank, 2018). Base year 2010.

A.1.5 Real* monthly per capita income, Peru 2004 – 2015

Year	Observations	Mean	Median	Minimum	Maximum	Gini
2004	19502	1277.81	721.38	0.00	116927.53	0.5805
2005	19895	1263.63	682.89	0.00	139560.55	0.5926
2006	20577	1416.62	786.02	0.00	79576.63	0.5878
2007	22204	1659.99	911.34	0.00	124390.45	0.5904
2008	21502	1740.60	984.61	0.00	131025.39	0.5803
2009	21753	1829.57	1027.13	0.00	93514.69	0.5786
2010	21496	1898.89	1128.03	0.00	111332.34	0.5583
2011	24809	1983.25	1216.17	0.00	121787.91	0.5560
2012	25091	2037.08	1293.44	0.00	172422.94	0.5402
2013	30453	2117.76	1318.58	0.00	111201.30	0.5407
2014	30848	2170.81	1326.66	0.00	270784.99	0.5438
2015	32188	2133.62	1309.35	0.00	96739.02	0.5392

Source: own estimations using microdata from ENAHO, 2004-2015.
*Real income estimation is based on Consumer Price Index (CPI) per year, published by the World Bank (World Bank, 2018). Base year 2010.

A.1.6 Real* monthly household per capita income, Uruguay 2006 – 2015

Year	Observations	Mean	Median	Minimum	Maximum	Gini
2006	85313	7966.30	5641.17	0.00	388996.88	0.4378
2007	49136	9863.86	6835.64	0.00	530080.64	0.4504
2008	50397	10730.86	7597.59	0.00	561433.97	0.4364
2009	46936	11834.86	8272.75	0.00	619690.27	0.4413
2010	46550	12143.67	8663.83	0.00	580500.00	0.4273
2011	46669	12425.91	9155.16	0.00	428714.64	0.4074
2012	43839	12867.71	9678.74	0.00	442173.83	0.3976
2013	46622	13041.58	9970.43	0.00	258436.09	0.3871
2014	48583	13719.67	10417.30	0.00	403885.29	0.3878
2015	45391	13941.49	10562.31	0.00	365792.64	0.3992

Source: own estimations using microdata from ECH, 2006-2015.
*Real income estimation is based on Consumer Price Index (CPI) per year, published by the World Bank (World Bank, 2018). Base year 2010.

A.2 Middle class income thresholds by country for relative and absolute definitions

A.2.1 Middle class income thresholds, relative and absolute definitions, Brazil 2004 – 2015

Relative definitions

Year	MC-1		MC-2		MC-3		R-1		R-2		R-3	
	Min	Max	Min	Max	Min	Max	Min	Max	Min	Max	Min	Max
2004	130.00	390.00	195.00	325.00	156.00	585.00	106.75	585.00	196.17	306.25	106.75	1000.00
2005	147.50	442.50	221.25	368.75	177.00	663.75	120.17	640.00	220.17	346.67	120.17	1100.00
2006	165.00	495.00	247.50	412.50	198.00	742.50	140.25	712.50	250.11	390.00	140.25	1200.00
2007	182.50	547.50	273.75	456.25	219.00	821.25	151.17	766.67	275.20	423.50	151.17	1300.00
2008	207.50	622.50	311.25	518.75	249.00	933.75	180.17	858.00	316.75	485.00	180.17	1461.67
2009	225.71	677.14	338.57	564.29	270.86	1015.71	199.11	931.25	348.40	520.00	199.11	1543.33
2011	272.50	817.50	408.75	681.25	327.00	1226.25	249.25	1100.00	433.50	637.50	249.25	1800.00
2012	311.00	933.00	466.50	777.50	373.20	1399.50	288.40	1250.00	500.22	725.00	288.40	2000.00
2013	339.00	1017.00	508.50	847.50	406.80	1525.50	313.33	1356.00	550.33	800.00	313.33	2225.00
2014	362.00	1086.00	543.00	905.00	434.40	1629.00	362.17	1500.00	625.00	890.00	362.17	2462.00
2015	394.00	1182.00	591.00	985.00	472.80	1773.00	375.20	1557.50	637.40	900.00	375.20	2500.00

Absolute definitions

Year	MC-4		MC-5		MC-6		MC-7	
	Min	Max	Min	Max	Min	Max	Min	Max
2004	2.04	13.24	10.19	50.94	2.04	10.19	12.23	50.94
2005	2.12	13.79	10.60	53.02	2.12	10.60	12.73	53.02
2006	2.20	14.28	10.99	54.93	2.20	10.99	13.18	54.93
2007	2.28	14.81	11.39	56.96	2.28	11.39	13.67	56.96
2008	2.43	15.80	12.15	60.77	2.43	12.15	14.58	60.77
2009	2.59	16.83	12.94	64.72	2.59	12.94	15.53	64.72
2011	2.94	19.12	14.71	73.55	2.94	14.71	17.65	73.55
2012	3.12	20.27	15.59	77.95	3.12	15.59	18.71	77.95
2013	3.30	21.44	16.49	82.46	3.30	16.49	19.79	82.46
2014	3.49	22.71	17.47	87.36	3.49	17.47	20.97	87.36
2015	3.72	24.17	18.59	92.97	3.72	18.59	22.31	92.97

Source: own estimations using microdata from PNAD, 2004-2015. Current values.

A.2.2 Middle class income thresholds, relative and absolute definitions, Chile 2000 – 2011

Relative definitions

Year	MC-1		MC-2		MC-3		R-1		R-2		R-3	
	Min	Max	Min	Max	Min	Max	Min	Max	Min	Max	Min	Max
2000	130.00	390.00	195.00	325.00	156.00	585.00	106.75	585.00	196.17	306.25	106.75	1000.00
2003	147.50	442.50	221.25	368.75	177.00	663.75	120.17	640.00	220.17	346.67	120.17	1100.00
2006	165.00	495.00	247.50	412.50	198.00	742.50	140.25	712.50	250.11	390.00	140.25	1200.00
2009	182.50	547.50	273.75	456.25	219.00	821.25	151.17	766.67	275.20	423.50	151.17	1300.00
2011	207.50	622.50	311.25	518.75	249.00	933.75	180.17	858.00	316.75	485.00	180.17	1461.67

Absolute definitions

Year	MC-4		MC-5		MC-6		MC-7	
	Min	Max	Min	Max	Min	Max	Min	Max
2000	2.04	13.24	10.19	50.94	2.04	10.19	12.23	50.94
2003	2.12	13.79	10.60	53.02	2.12	10.60	12.73	53.02
2006	2.20	14.28	10.99	54.93	2.20	10.99	13.18	54.93
2009	2.28	14.81	11.39	56.96	2.28	11.39	13.67	56.96
2011	2.43	15.80	12.15	60.77	2.43	12.15	14.58	60.77

Source: Own estimations using microdata from CASEN, 2000-2011. Current values.

A.2.3 Middle class income thresholds, relative and absolute definitions, Mexico 2000 – 2014

Relative definitions

Year	MC-1		MC-2		MC-3		R-1		R-2		R-3	
	Min	Max	Min	Max	Min	Max	Min	Max	Min	Max	Min	Max
2000	581.97	2613.53	977.90	1534.43	581.97	4042.77	616.77	1850.31	925.16	1541.93	740.12	2775.47
2002	686.29	2963.66	1139.94	1744.44	686.29	4566.73	703.39	2110.16	1055.08	1758.47	844.07	3165.25
2004	904.50	4112.69	1509.23	2351.37	904.50	6510.82	939.04	2817.12	1408.56	2347.60	1126.85	4225.68
2005	850.64	4071.07	1444.65	2293.70	850.64	6261.36	907.38	2722.15	1361.08	2268.46	1088.86	4083.23
2006	998.13	4481.27	1655.28	2574.28	998.13	7049.02	1028.37	3085.10	1542.55	2570.92	1234.04	4627.65
2008	1129.13	4961.20	1847.19	2858.81	1129.13	7631.52	1147.57	3442.71	1721.35	2868.92	1377.08	5164.06
2010	1119.71	4752.83	1812.75	2772.57	1119.71	7329.33	1121.92	3365.75	1682.88	2804.79	1346.30	5048.63
2012	1207.14	5086.96	1910.63	2932.33	1207.14	7829.42	1169.70	3509.11	1754.55	2924.25	1403.64	5263.66
2014	1390.25	5302.45	2138.41	3178.76	1390.25	8174.34	1292.23	3876.69	1938.35	3230.58	1550.68	5815.04

Absolute definitions

Year	MC-4		MC-5		MC-6		MC-7	
	Min	Max	Min	Max	Min	Max	Min	Max
2000	12.19	79.22	60.94	304.68	12.19	60.94	73.12	304.68
2002	13.11	85.20	65.54	327.68	13.11	65.54	78.64	327.68
2004	14.34	93.22	71.71	358.53	14.34	71.71	86.05	358.53
2005	14.25	92.65	71.27	356.34	14.25	71.27	85.52	356.34
2006	14.37	93.42	71.86	359.31	14.37	71.86	86.23	359.31
2008	14.94	97.10	74.70	373.48	14.94	74.70	89.63	373.48
2010	15.34	99.68	76.68	383.39	15.34	76.68	92.01	383.39
2012	15.72	102.16	78.59	392.94	15.72	78.59	94.30	392.94
2014	16.09	104.59	80.45	402.27	16.09	80.45	96.54	402.27

Source: own estimations using microdata from ENIGH, 2000-2014. Current values.

A.2.4 Middle class income thresholds, relative and absolute definitions, Paraguay 2002 – 2015

Relative definitions

Year	MC-1		MC-2		MC-3		R-1		R-2		R-3	
	Min	Max	Min	Max	Min	Max	Min	Max	Min	Max	Min	Max
2002	108092.00	619305.38	204610.00	348792.00	108092.00	967799.69	133468.20	400404.50	200202.30	333670.40	160161.80	600606.80
2003	135239.70	665000.00	235183.83	374417.50	135239.70	1050489.80	147998.20	443994.50	221997.30	369995.40	177597.80	665991.80
2004	155771.67	710000.00	258823.41	402460.00	155771.67	1087601.40	162240.60	486721.90	243360.90	405601.60	194688.80	730082.80
2005	187078.00	900000.00	317079.59	506529.34	187078.00	1397385.60	199158.10	597474.30	298737.10	497895.30	238989.70	896211.40
2006	185665.22	869333.25	324000.00	501488.50	185665.22	1407511.60	200595.50	601786.50	300893.30	501488.80	240714.60	902679.80
2007	222711.00	1050176.40	379717.00	593621.50	222711.00	1610086.40	238180.00	714540.10	357270.10	595450.10	285816.00	1071810.00
2008	250929.41	1157652.00	422787.00	665875.63	250929.41	1830155.00	262666.90	788000.60	394000.30	656667.20	315200.30	1182001.00
2009	241746.33	1225664.50	436388.84	706675.00	241746.33	1862771.00	276666.50	829999.50	414999.80	691666.30	331999.80	1244999.00
2010	298703.50	1394444.40	520966.25	815000.00	298703.50	2130555.30	328573.50	985720.50	492860.30	821433.80	394288.20	1478581.00
2011	339693.41	1547111.10	598322.50	915891.75	339693.41	2295375.00	371271.10	1113813.00	556906.60	928177.60	445525.30	1670720.00
2012	386646.00	1678549.00	670689.75	1013160.30	386646.00	2487421.30	412408.30	1237225.00	618612.50	1031021.00	494890.00	1855838.00
2013	443168.00	1915978.60	750376.00	1158326.60	443168.00	2888433.00	467851.60	1403555.00	701777.40	1169629.00	561421.90	2105332.00
2014	487949.41	2000964.60	781013.75	1202643.00	487949.41	2912000.00	485400.50	1456202.00	728100.80	1213501.00	582480.60	2184302.00
2015	464154.34	2020000.00	784824.50	1201805.00	464154.34	2930720.00	487332.60	1461998.00	730998.90	1218332.00	584799.10	2192997.00

Absolute definitions

Year	MC-4		MC-5		MC-6		MC-7	
	Min	Max	Min	Max	Min	Max	Min	Max
2002	2746.18	17850.15	13730.89	68654.43	2746.18	13730.89	16477.06	68654.43
2003	3021.27	19638.23	15106.33	75531.66	3021.27	15106.33	18127.60	75531.66
2004	3204.64	20830.15	16023.19	80115.95	3204.64	16023.19	19227.83	80115.95
2005	3417.55	22214.08	17087.75	85438.77	3417.55	17087.75	20505.30	85438.77
2006	3517.37	22862.88	17586.83	87934.16	3517.37	17586.83	21104.20	87934.16
2007	3760.79	24445.12	18803.94	94019.70	3760.79	18803.94	22564.73	94019.70
2008	4032.78	26213.05	20163.88	100819.40	4032.78	20163.88	24196.66	100819.40
2009	4084.16	26547.01	20420.78	102103.90	4084.16	20420.78	24504.93	102103.90
2010	4281.01	27826.59	21405.07	107025.30	4281.01	21405.07	25686.08	107025.30
2011	4454.68	28955.42	22273.40	111367.00	4454.68	22273.40	26728.08	111367.00
2012	4581.76	29781.45	22908.81	114544.00	4581.76	22908.81	27490.57	114544.00
2013	4546.87	29554.66	22734.35	113671.80	4546.87	22734.35	27281.22	113671.80
2014	4696.25	30525.64	23481.27	117406.30	4696.25	23481.27	28177.52	117406.30
2015	4649.93	30224.52	23249.63	116248.10	4649.93	23249.63	27899.55	116248.10

Source: own estimations using microdata from EPH, 2002-2015. Current values.

A.2.5 Middle class income thresholds, relative and absolute definitions, Peru 2004 – 2015

Relative definitions

Year	MC-1		MC-2		MC-3		R-1		R-2		R-3	
	Min	Max	Min	Max	Min	Max	Min	Max	Min	Max	Min	Max
2004	191.34	1512.96	451.44	822.04	191.34	2345.62	309.26	927.77	463.88	773.14	371.11	1391.65
2005	179.22	1510.22	421.79	807.71	179.22	2383.41	297.48	892.45	446.23	743.71	356.98	1338.68
2006	204.03	1731.44	502.32	938.82	204.03	2716.18	349.27	1047.80	523.90	873.16	419.12	1571.69
2007	245.77	2040.94	596.52	1100.04	245.77	3258.78	412.16	1236.48	618.24	1030.40	494.59	1854.72
2008	284.81	2277.42	690.80	1256.58	284.81	3561.00	471.07	1413.20	706.60	1177.66	565.28	2119.79
2009	328.40	2443.67	752.51	1339.93	328.40	3730.57	505.83	1517.50	758.75	1264.58	607.00	2276.25
2010	381.63	2617.83	853.73	1479.17	381.63	3962.00	564.01	1692.04	846.02	1410.04	676.82	2538.06
2011	411.83	2774.33	945.50	1617.80	411.83	4186.67	628.58	1885.74	942.87	1571.45	754.30	2828.61
2012	475.17	3012.00	1048.33	1773.44	475.17	4511.17	692.94	2078.83	1039.42	1732.36	831.53	3118.25
2013	509.01	3259.11	1098.49	1878.75	509.01	4913.83	726.31	2178.93	1089.46	1815.77	871.57	3268.39
2014	550.75	3414.50	1148.67	1950.22	550.75	5193.00	754.34	2263.01	1131.50	1885.84	905.20	3394.51
2015	580.33	3488.33	1177.33	1983.00	580.33	5279.80	770.96	2312.89	1156.45	1927.41	925.16	3469.33

Absolute definitions

Year	MC-4		MC-5		MC-6		MC-7	
	Min	Max	Min	Max	Min	Max	Min	Max
2004	2.68	17.41	13.39	66.96	2.68	13.39	16.07	66.96
2005	2.69	17.46	13.43	67.14	2.69	13.43	16.11	67.14
2006	2.81	18.23	14.03	70.13	2.81	14.03	16.83	70.13
2007	2.77	18.03	13.87	69.33	2.77	13.87	16.64	69.33
2008	2.75	17.87	13.75	68.75	2.75	13.75	16.50	68.75
2009	2.79	18.11	13.93	69.64	2.79	13.93	16.71	69.64
2010	2.91	18.91	14.55	72.73	2.91	14.55	17.45	72.73
2011	3.04	19.78	15.21	76.06	3.04	15.21	18.25	76.06
2012	3.02	19.65	15.12	75.58	3.02	15.12	18.14	75.58
2013	3.01	19.55	15.04	75.18	3.01	15.04	18.04	75.18
2014	3.03	19.70	15.15	75.75	3.03	15.15	18.18	75.75
2015	3.08	19.99	15.38	76.90	3.08	15.38	18.46	76.90

Source: own estimations using microdata from ENAHO, 2004-2015. Current values.

A.2.6 Middle class income thresholds, relative and absolute definitions, Uruguay 2006 – 2015

Relative definitions

Year	MC-1		MC-2		MC-3		R-1		R-2		R-3	
	Min	Max	Min	Max	Min	Max	Min	Max	Min	Max	Min	Max
2006	2250.04	8030.95	3510.00	5101.33	2250.04	11618.50	2117.08	6351.25	3175.62	5292.71	2540.50	9526.87
2007	2881.95	10845.00	4568.75	6739.67	2881.95	16042.22	2773.52	8320.55	4160.27	6933.79	3328.22	12480.82
2008	3600.75	12684.00	5542.00	8000.00	3600.75	18286.67	3325.00	9975.00	4987.50	8312.50	3990.00	14962.50
2009	4191.38	15000.00	6409.25	9420.38	4191.38	21983.50	3877.50	11632.50	5816.25	9693.75	4653.00	17448.75
2010	4800.50	16497.11	7201.00	10453.67	4800.50	23731.50	4331.92	12995.75	6497.88	10829.79	5198.30	19493.63
2011	5600.00	18181.42	8318.52	11768.57	5600.00	25700.13	4948.08	14844.24	7422.12	12370.20	5937.70	22266.37
2012	6463.13	20615.64	9508.80	13499.67	6463.13	28713.94	5654.66	16963.97	8481.99	14136.64	6785.59	25445.96
2013	7372.00	22607.00	10749.36	14950.17	7372.00	31349.35	6324.58	18973.74	9486.87	15811.45	7589.50	28460.62
2014	8515.53	25659.66	12233.00	16962.00	8515.53	35291.00	7194.67	21584.00	10792.00	17986.67	8633.60	32376.00
2015	9214.70	28304.00	13439.33	18739.89	9214.70	39518.83	7927.00	23781.00	11890.50	19817.50	9512.40	35671.50

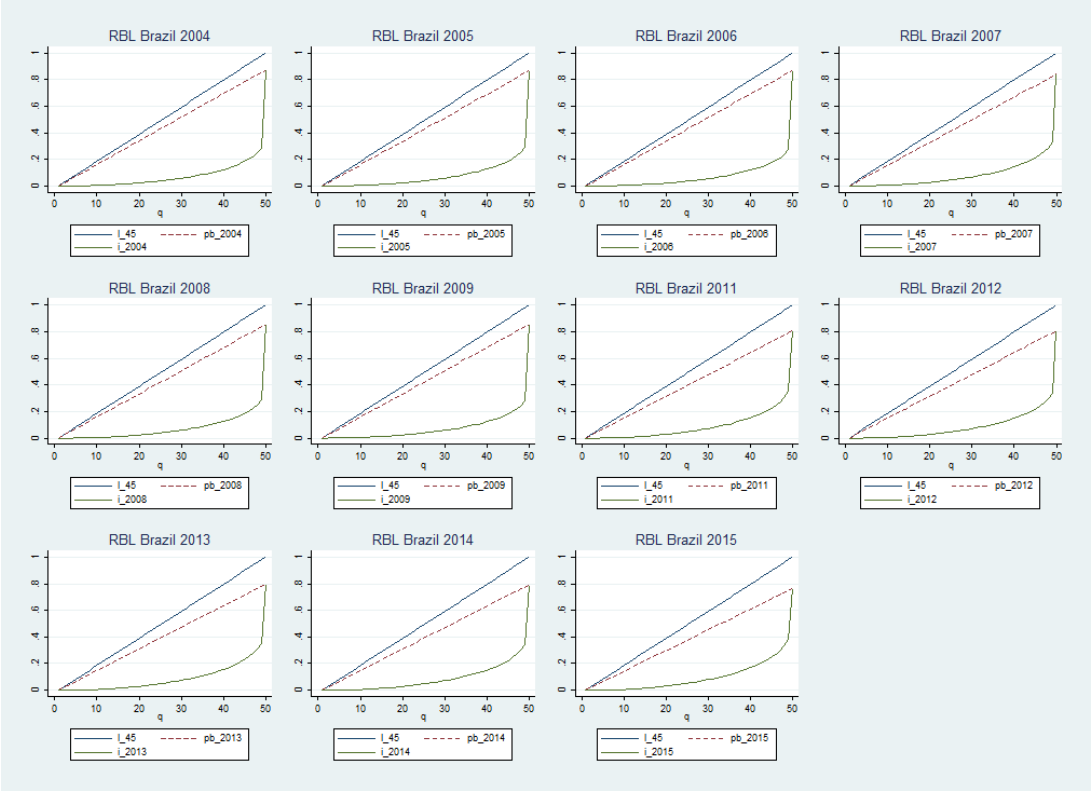
Absolute definitions

Year	MC-4		MC-5		MC-6		MC-7	
	Min	Max	Min	Max	Min	Max	Min	Max
2006	22.87	148.63	114.33	571.66	22.87	114.33	137.20	571.66
2007	24.37	158.41	121.86	609.28	24.37	121.86	146.23	609.28
2008	25.82	167.83	129.10	645.50	25.82	129.10	154.92	645.50
2009	27.61	179.47	138.05	690.27	27.61	138.05	165.66	690.27
2010	28.62	186.01	143.08	715.40	28.62	143.08	171.70	715.40
2011	30.56	198.66	152.82	764.08	30.56	152.82	183.38	764.08
2012	32.58	211.76	162.89	814.47	32.58	162.89	195.47	814.47
2013	34.68	225.39	173.38	866.88	34.68	173.38	208.05	866.88
2014	37.26	242.17	186.28	931.41	37.26	186.28	223.54	931.41
2015	40.18	261.17	200.90	1004.49	40.18	200.90	241.08	1004.49

Source: own estimations using microdata from ECH, 2006-2015. Current values.

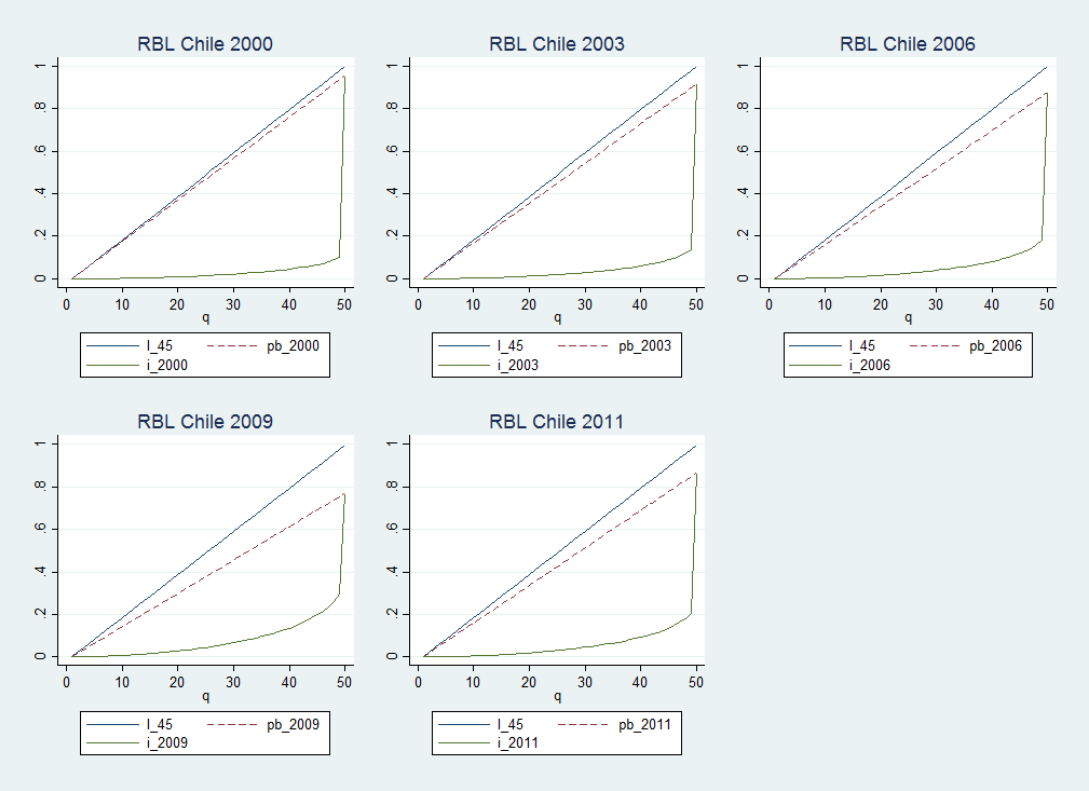
A.3 Relative Bipolarization Lorenz curves for all countries and years

A.3.1 Relative Bipolarization Lorenz curves, Brazil 2004 – 2015



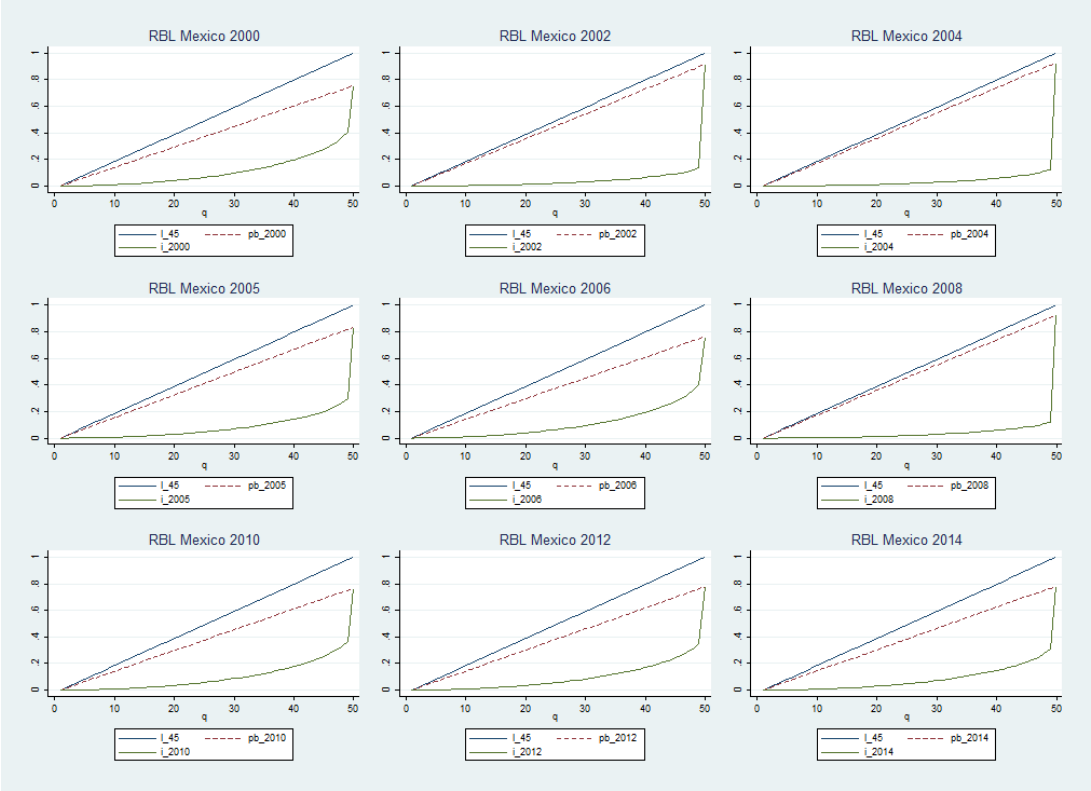
Source: own estimations using microdata from PNAD 2004-2015.

A.3.2 Relative Bipolarization Lorenz curves, Chile 2000 – 2011



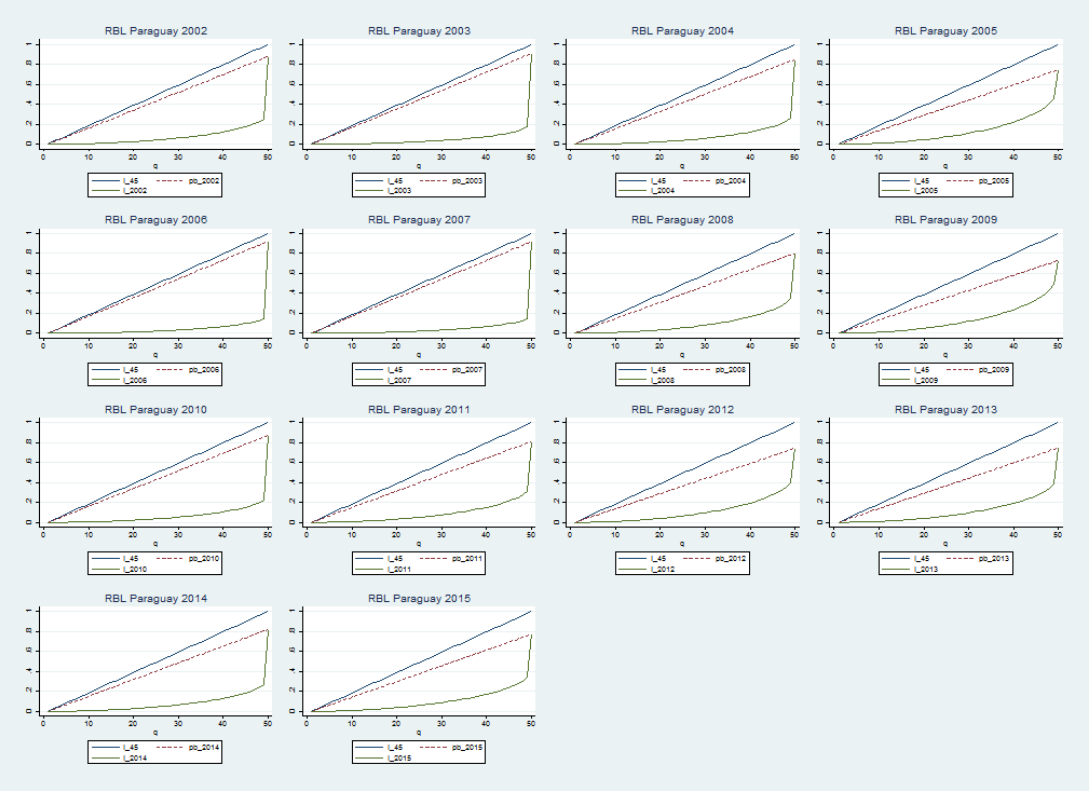
Source: own estimations using microdata from CASEN, 2000-2011

A.3.3 Relative Bipolarization Lorenz curves, Mexico 2000 – 2014



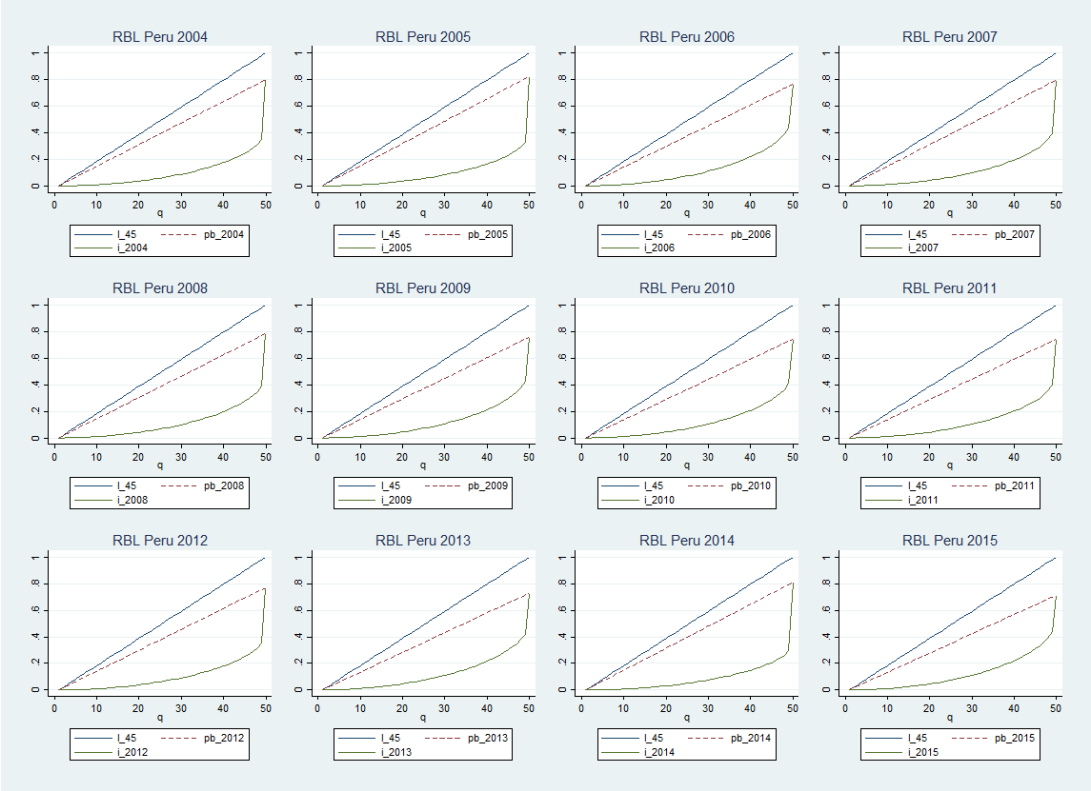
Source: own estimations using microdata from ENIGH, 2000-2014.

A.3.4 Relative Bipolarization Lorenz curves, Paraguay 2002 – 2015



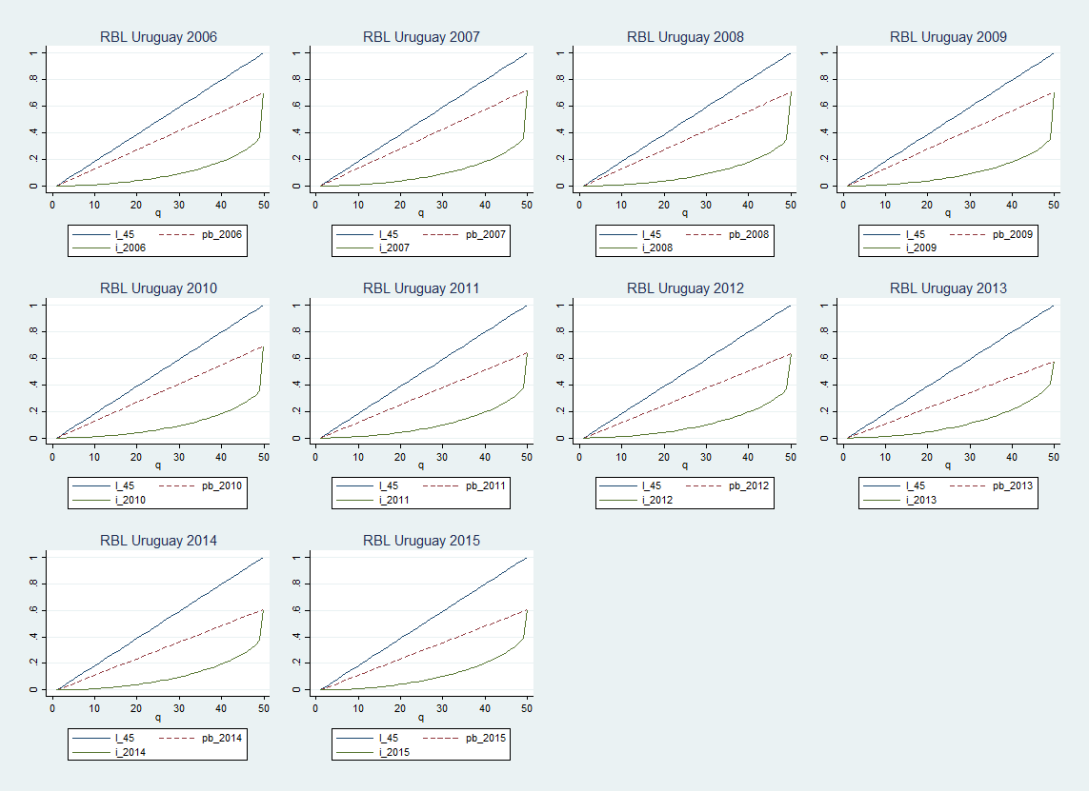
Source: own estimations using microdata from EPH, 2002-2015.

A.3.5 Relative Bipolarization Lorenz curves, Peru 2004 – 2015



Source: own estimations using microdata from ENAHO, 2004-2015.

A.3.6 Relative Bipolarization Lorenz curves, Uruguay 2006 – 2015



Source: own estimations using microdata from ECH, 2006-2015.

Appendix B
Multidimensional middle class in Mexico

B.1 Class dimensions and their wellbeing domains

Dimension	Wellbeing domain	Description
Education	Technical or professional education	Achievement or studying any level of technical or professional education
Health	SP access or no access to health services	Access to health services provided by SP, or no access to health services of any kind, controlled by the use of *Oportunidades* social program and healers
	Access to SS health services or private GP	Access to any type of SS health services and common use of private GP practices
	Private medical insurance	Full coverage by private medical insurance
Dwelling characteristics	Overcrowding	Ratio of total residents by total rooms available in the dwelling for sleep
	Lightbulbs	Ratio of total lightbulbs by total rooms in the dwelling
Household assets	Landline and mobile phones	Household per capita landline and mobile phones
	Cars, vans or pick-ups	Household per capita cars, vans or pick-ups, controlled by number of new cars
	Hi-Fi, radios and music recorders	Household per capita Hi-Fi, radios and music recorders
	TV	Household per capita TVs
	VCR and DVD	Household per capita VCRs and DVDs
	Computer and printer	Household per capita computers and printers
Occupation characteristics	Low skilled occupations	Operators, peasants, drivers, domestic work, lower rank soldiers
	Skilled occupations	Administrative, managerial and skilled occupations, upper rank soldiers, retired and students
	High profile occupations	Directors, high profile politicians and civil servants, military commanders and chiefs, landlords

B.2 Estimation of multidimensional poverty in Mexico, 2012, areas in Figure II.1

Area	Percentage	Population (millions)
A + B	45.5	53.3
A	9.8	11.5
C	28.6	33.5
D	6.2	7.2
E	19.8	23.2

Source: CONEVAL (2018)

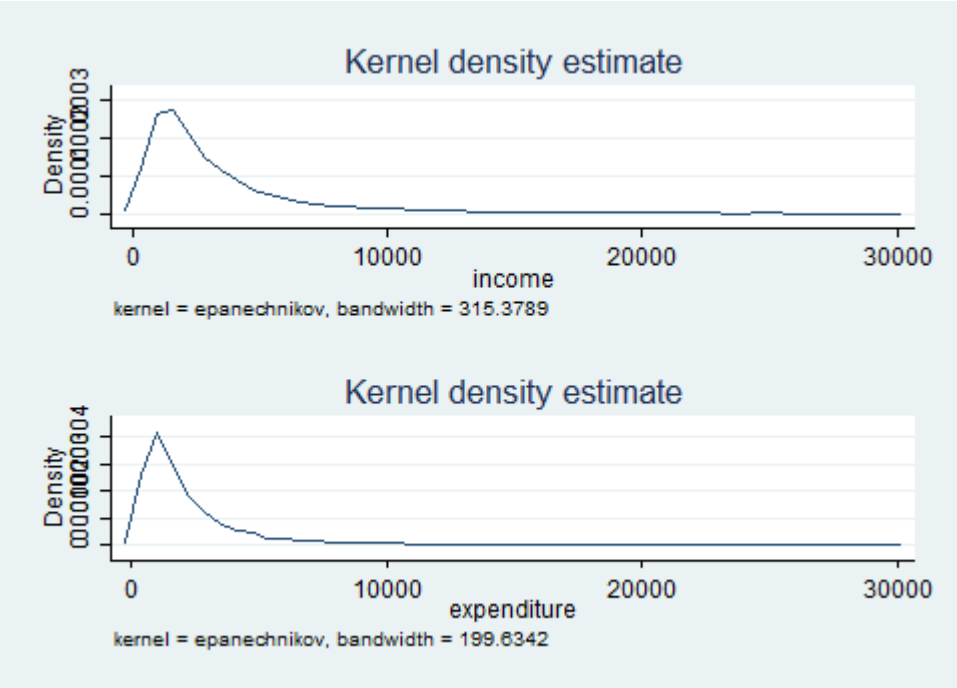
B.3 Descriptive statistics and density functions for income and expenditure

B.3.1 Descriptive statistics, household per capita income and expenditure

Variable	Mean	Median	Minimum	Maximum
Per capita income	3,818.36	2,268.65	48.91	264,298.10
Per capita expenditure	2,519.20	1,563.48	0.00	105,645.60

Source: own estimations using microdata from ENIGH, 2012.

B.3.2 Density functions, household per capita income and expenditure



Source: own estimations using microdata from ENIGH, 2012.

B.4 Descriptive analysis for the middle class counting approach dimensions and wellbeing domains

B.4.1 Descriptive statistics, wellbeing domains by middle class dimensions

Wellbeing domain	Observations	Mean	Standard deviation	Minimum	Maximum
Eduaction					
HH education ratio	211,595	0.113	0.222	0.0	2.0
Health					
Deprivation health	211,595	0.573	0.432	0.0	1.0
Mid-health	211,595	0.419	0.427	0.0	1.0
Affluent health	211,595	0.008	0.065	0.0	1.0
Dwelling characteristics					
Overcrowding index	211,595	2.494	1.407	0.1	17.0
Lightbulb index	209,959	1.830	0.912	0.1	21.5
Household assets					
Automobil level	211,595	1.382	0.538	1.0	3.0
Audio level	211,595	1.538	0.506	1.0	3.0
TV level	211,595	1.476	0.512	1.0	3.0
Video level	211,595	1.496	0.508	1.0	3.0
Computer level	211,595	1.291	0.459	1.0	3.0
Occupation and employment					
Deprived worker	211,595	1.226	1.189	0.0	10.0
Mid-worker	211,595	0.660	0.881	0.0	7.0
Affluent worker	211,595	0.032	0.188	0.0	3.0
Unemployed	211,595	0.114	0.369	0.0	5.0
Landlord	211,595	0.003	0.059	0.0	3.0
Retired	211,595	0.079	0.294	0.0	3.0
Housekeeping	211,595	0.763	0.770	0.0	7.0
Student	211,595	0.579	0.823	0.0	7.0
Disabled	211,595	0.057	0.254	0.0	4.0

Source: own estimations using microdata from MCS-ENIGH, 2012.

B.4.2 Correlation analysis, dimensions of the middle class

	Eduaction	Health	Dwelling characteristics	Household assets	Occupation
Eduaction	1.000				
Health	0.365	1.000			
Dwelling characteristics	0.245	0.250	1.000		
Household assets	0.363	0.346	0.344	1.000	
Occupation	0.443	0.388	0.240	0.347	1.000

Source: own estimations using microdata from MCS-ENIGH 2012.

Appendix C
Identification of the middle class using predicted labour income

C.1 Variables used in the predicted income estimation

Variable	MCS-ENIGH 2012			ENOE 2005			ENOE 2011		
	Obs	Mean	Standard deviation	Obs	Mean	Standard deviation	Obs	Mean	Standard deviation
Household per capita income	56,424	1,853	3,798	409,319	1,734	3,013	418,502	1,513	2,289
Household head male	56,424	0.75	0.43	409,319	0.77	0.42	418,502	0.74	0.44
Household head age	56,424	48.62	15.81	409,319	47.18	15.41	418,502	48.47	15.78
Household head married	56,424	1.58	1.07	409,319	1.61	1.01	418,502	1.56	1.06
Household head years education	56,424	7.99	4.97	409,319	8.00	5.15	418,502	8.57	5.04
Household dependency rate	56,424	0.57	0.66	409,319	0.61	0.68	418,502	0.54	0.64
Average household workers	56,424	0.47	0.29	409,319	0.44	0.28	418,502	0.45	0.29
Urban locality	56,424	0.75	0.43	409,319	0.84	0.37	418,502	0.83	0.38

Source: own estimations using microdata from MCS-ENIGH 2012, ENOE 2005 and ENOE 2011.

C.2 Econometric estimation of the predicted per capita labour income

Variable	MCS-ENIGH	ENOE 2005	ENOE 2011
Household per capita income	46.406***	204.34***	180.81***
	(16.775)	(12.661)	(8.3534)
Household head male	-8.2737***	5.5046***	1.0695***
	(0.4544)	(0.3046)	(0.2127)
Household head age	93.559***	-50.529***	-65.204***
	(6.8897)	(5.2584)	(3.4778)
Household head married	47.774***	151.32***	105.43***
	(1.4848)	(0.9315)	(0.6865)
Household head years education	25.950**	-30.425***	-6.8094
	(11.184)	(6.8699)	(5.2052)
Household dependency rate	1888.9***	3544.9***	3136.9***
	(23.488)	(16.505)	(11.254)
Average household workers	204.97***	284.86***	236.40***
	(19.083)	(12.077)	(8.5601)
Urban locality	407.93***	-1724.0***	-1343.6***
	(51.020)	(35.020)	(24.833)
Number of observations	22,915	409,319	418,502
R^2	0.3242	0.2068	0.2619

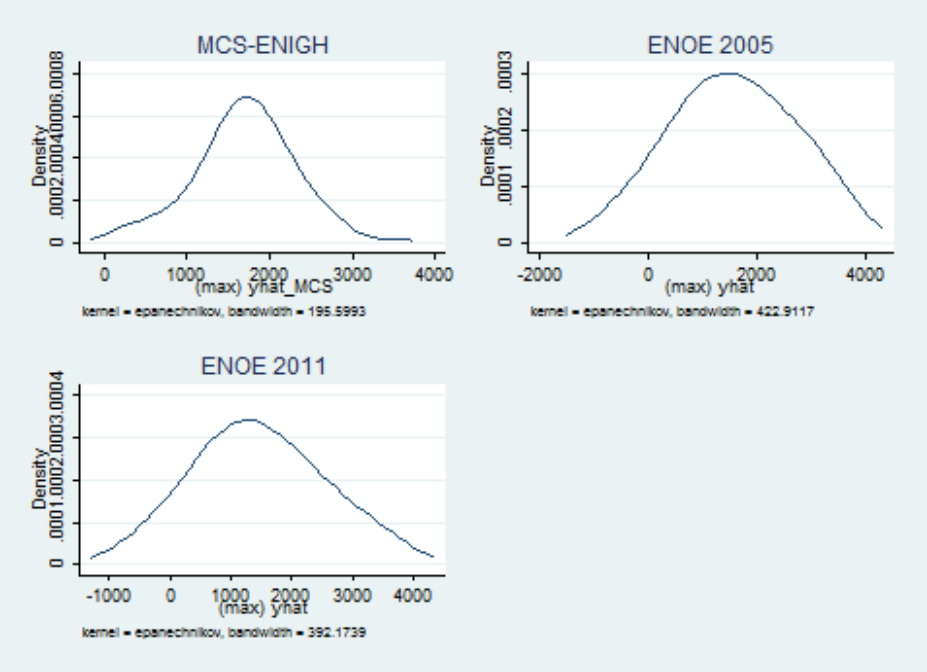
Source: own estimations using microdata from MCS-ENIGH 2012, ENOE 2005 and ENOE 2011.

C.3 Predicted per capita household labour income, descriptive statistics

Sample	Observations	Mean	Standard deviation
MCS-ENIGH	22,915	1,680.43	622.84
ENOE 2005	409,319	1,734.15	1,369.94
ENOE 2011	418,502	1,513.33	1,171.28

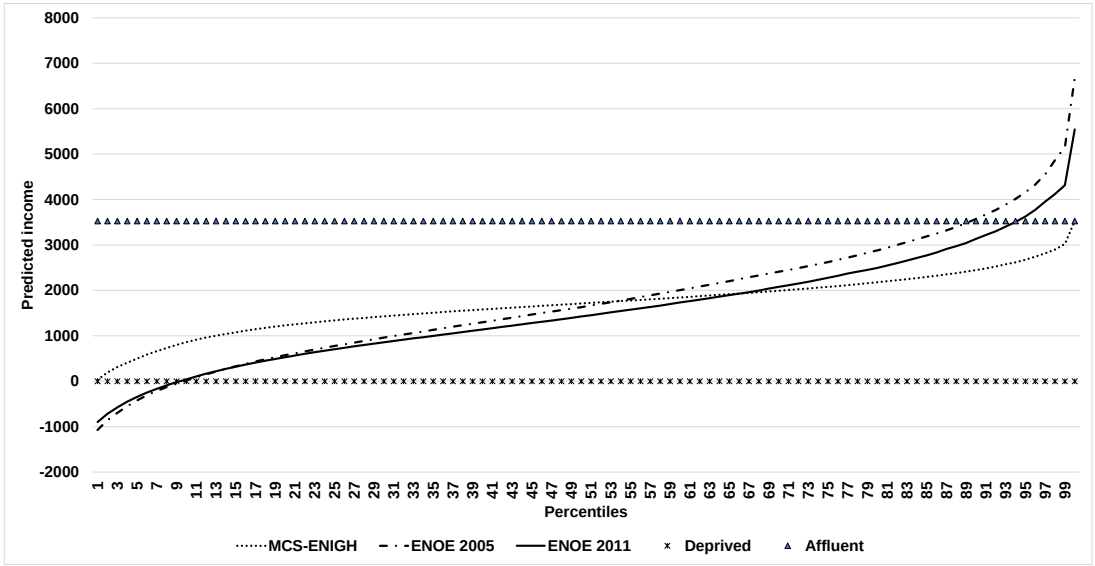
Source: own estimations using microdata from MCS-ENIGH 2012, ENOE 2005 and ENOE 2011.

C.4 Predicted per capita household labour income, densities



Source: own estimations using microdata from MCS-ENIGH 2012, ENOE 2005 and ENOE 2011.

C.5 Percentile distribution of the predicted household per capita labour income by sample survey, deprivation and affluent thresholds



Source: own estimations using microdata from MCS-ENIGH 2012, ENOE 2005 and ENOE 2011.

C.6 Placebo tests, 2005-2007 and 2015-1017

Women labour market participation 2005-2007

Interaction effect BS	Interaction effect RSE
0.0173***	0.0173***
(0.0062)	(0.0063)

Source: own estimations using microdata from ENOE, 2005 and 2007.
Note: Standard errors in brackets; levels of significance at 90%, 95% and 99% as *, ** and *** respectively.

Women participation in informal labour market 2005-2007

Interaction effect BS	Interaction effect BS^C	Interaction effect RSE	Interaction effect RSE^C
-0.0116	-0.0028	-0.0115	-0.0028
(0.0098)	(0.0087)	(0.0097)	(0.0028)

Source: own estimations using microdata from ENOE, 2005 and 2007.
Note: Standard errors in brackets; levels of significance at 90%, 95% and 99% as *, ** and *** respectively.

Women labour market participation, hours 2005-2007

Interaction effect BS	Interaction effect BS^C	Interaction effect RSE	Interaction effect RSE^C
0.0086	0.0004	0.0086	0.0044
(0.0142)	(0.0128)	(0.0129)	(0.0126)

Source: own estimations using microdata from ENOE, 2005 and 2007.
Note: Standard errors in brackets; levels of significance at 90%, 95% and 99% as *, ** and *** respectively.

Women labour market participation 2015-2017

Interaction effect BS	Interaction effect RSE
-0.0229***	-0.0229***
(0.0056)	(0.0057)

Source: own estimations using microdata from ENOE, 2015 and 2017.
Note: Standard errors in brackets; levels of significance at 90%, 95% and 99% as *, ** and *** respectively.

Women participation in informal labour market 2015-2017

Interaction effect BS	Interaction effect BS^C	Interaction effect RSE	Interaction effect RSE^C
-0.0189	-0.0085	-0.0189	-0.0085
(0.0085)	(0.0058)	(0.0078)	(0.0069)

Source: own estimations using microdata from ENOE, 2015 and 2017.
Note: Standard errors in brackets; levels of significance at 90%, 95% and 99% as *, ** and *** respectively.

Women labour market participation, hours 2015-2017

Interaction effect BS	Interaction effect BS^C	Interaction effect RSE	Interaction effect RSE^C
-0.0125	-0.0204	-0.0125	-0.0204
(0.0099)	(0.0097)	(0.0102)	(0.0100)

Source: own estimations using microdata from ENOE, 2015 and 2017.
Note: Standard errors in brackets; levels of significance at 90%, 95% and 99% as *, ** and *** respectively.

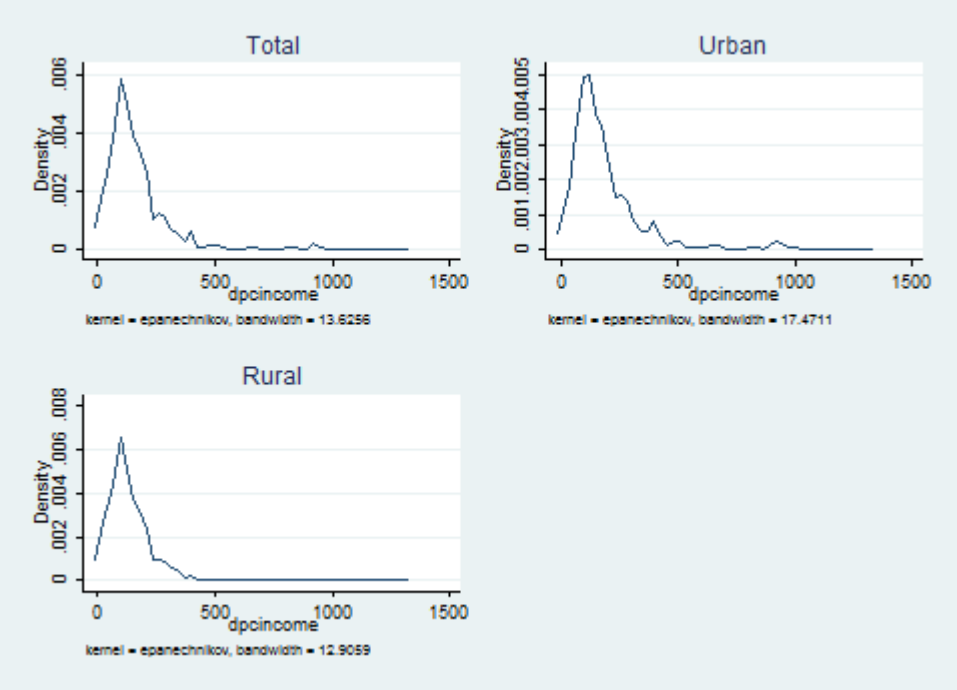
Appendix D
Financial exclusion analysis

D.1 Daily per capita income by urbanization context in Mexican pesos, descriptive statistics

Urbanization context	Obs	Mean	Standard deviation	Minimum	Maximum
Urban	2,722	188.25	171.14	0.03	1578.95
Rural	2,669	136.94	125.29	0.03	2335.54
National	**5,391**	**162.75**	**152.28**	**0.03**	**2335.54**

Source: own estimations using microdata from ENCAS, 2011.

D.2 Daily per capita income by urbanization context in Mexican pesos, distribution densities



Source: own estimations using microdata from ENCAS, 2011.

D.3 Multiple correspondence analysis for the household assets index

Number of obs	=	5,314
Total inertia	=	0.076
Number of axes	=	2

Categories	overall			dimension_1			dimension_2		
	mass	quality	%inert	coord	sqcorr	contrib	coord	sqcorr	contrib
Mobile phone									
0	0.062	0.993	0.057	0.98	0.991	0.06	0.621	0.001	0.024
1	0.063	0.993	0.056	-0.969	0.991	0.059	-0.614	0.001	0.024
Landline									
0	0.087	0.98	0.032	0.617	0.977	0.033	0.663	0.003	0.038
1	0.038	0.98	0.074	-1.425	0.977	0.077	-1.533	0.003	0.089
Fridge									
0	0.022	0.904	0.134	2.416	0.901	0.129	-2.154	0.002	0.102
1	0.103	0.904	0.029	-0.518	0.901	0.028	0.462	0.002	0.022
Gas stove									
0	0.014	0.9	0.12	2.873	0.895	0.115	-3.932	0.005	0.215
1	0.111	0.9	0.015	-0.359	0.895	0.014	0.491	0.005	0.027
TV									
0	0.006	0.992	0.049	2.935	0.978	0.051	-6.614	0.015	0.261
1	0.119	0.992	0.002	-0.147	0.978	0.003	0.332	0.015	0.013
VCR/DVD									
0	0.07	0.975	0.05	0.863	0.974	0.052	0.512	0.001	0.018
1	0.055	0.975	0.065	-1.106	0.974	0.067	-0.656	0.001	0.024
Washing machine									
0	0.053	0.914	0.115	1.459	0.914	0.112	0.363	0	0.007
1	0.072	0.914	0.083	-1.057	0.914	0.081	-0.263	0	0.005
Car/van/pickup									
0	0.085	0.96	0.038	0.677	0.957	0.039	0.707	0.003	0.042
1	0.04	0.96	0.08	-1.419	0.957	0.081	-1.482	0.003	0.089

Source: own estimations using microdata from ENCAS, 2011.

D.4 Propensity score estimations by social class

Variable	Total	Middle class	Deprived	Affluent
Male	0.3342*	0.4049*	2.4944	0.7637
	(0.1800)	(0.2127)	(1.5216)	(0.8530)
Age	-0.0077	-0.0012	-0.2552	0.0401
	(0.0261)	(0.0292)	(0.1859)	(0.1364)
Age square	0.0000	-0.0000	0.0016	-0.0004
	(0.0002)	(0.0003)	(0.0017)	(0.0015)
Years of education	-0.0200	-0.0179	-0.2765	0.0314
	(0.0148)	(0.0173)	(0.2039)	(0.0674)
HH head	-0.2986	-0.4668**	2.4240*	-1.5421
	(0.1835)	(0.2160)	(1.2655)	(0.9367)
Working	-0.0594	0.1091	-3.3420*	-1.1063
	(0.1616)	(0.1864)	(1.7735)	(0.9033)
HH total dependency rate	0.1871*	0.1689	3.0802	0.1461
	(0.1088)	(0.1216)	(2.8527)	(0.5403)
HH asset index	-0.9012	-0.9166	9.7227*	-6.5186
	(0.5655)	(0.6714)	(5.2983)	(4.9692)
Bonding networks	0.1787	0.2580	2.4769	3.3859
	(0.4674)	(0.5610)	(3.1710)	(2.6384)
Linking networks	-0.4423	-0.5539	-1.3597	0.5224
	(0.3274)	(0.3925)	(1.5365)	(1.2171)
Homicide rate[1]	0.0166***	0.0129**	5.6894	0.0331
	(0.0058)	(0.0062)	(11.480)	(0.0205)
Region 1	-0.1339	0.0446	-6.3653	-1.9244**
	(0.2004)	(0.2327)	(8.7934)	(0.8277)
Region 2	0.2424	0.2837	1.2441	-0.9694
	(0.1702)	(0.1928)	(1.3333)	(0.8087)
Amount borrowed	-0.0000	0.0000	-0.0000	-0.0000
	(0.0000)	(0.0000)	(0.0001)	(0.0000)
Constant	1.8699***	1.5741*	2.0958	4.5588
	(0.7201)	(0.8512)	(3.9757)	(3.4776)
N	489	368	65	56
Pseudo R2	0.0614	0.0594	0.5902	0.3713

Source: own estimations using microdata from ENCAS, 2011.
Note: Standard errors in brackets; levels of significance at 90%, 95% and 99% as *, ** and *** respectively.
1 Ratio of yearly drug-trafficking related homicides in the municipality per 100,000 inhabitants.