Public Disclosure Authorized

Public Disclosure Authorized

WPS5902

Policy Research Working Paper 5902

A Vulnerability Approach to the Definition of the Middle Class

Luis F. Lopez-Calva
Eduardo Ortiz-Juarez

The World Bank
Latin America and the Caribbean Region
Poverty, Equity and Gender Unit
December 2011



Policy Research Working Paper 5902

Abstract

Measurement of the middle class has recently come to the center of policy debate in middle-income countries as they search for the potential engines of growth and good governance. This debate assumes, first, that there is a meaningful definition of class, and second, that thresholds that define relatively homogeneous groups in terms of pre-determined sociological characteristics can be found empirically. This paper aims at proposing a view of the middle class based on vulnerability to poverty. Following this approach the paper exploits panel data to determine the amount of comparable income—associated with a low probability of falling into poverty—which could define the lower bound of the middle class. The paper looks at absolute thresholds, challenging the view that people above the poverty line are actually part of the middle class. The estimated lower threshold is used in cross-section surveys to quantify the size and the evolution of middle classes in Chile, Mexico, and Peru over the past two decades. The first relevant feature relates to the fact that the proposed thresholds lie around the 60th percentile of the distribution. The evidence also shows that the middle class has increased significantly in all three countries, suggesting that a higher number of households face lower probabilities of falling into poverty than they did in the 1990s. There is an important group of people, however, which cannot be defined as middle class from this perspective, but are not eligible for poverty programs according to traditional definitions of poverty.

This paper is a product of the Poverty, Equity and Gender Unit, Latin America and the Caribbean Region. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://econ.worldbank.org.
The author may be contacted at lflopezcalva@worldbank.org.

The Policy Research Working Paper Series disseminates the findings of work in progress to encourage the exchange of ideas about development issues. An objective of the series is to get the findings out quickly, even if the presentations are less than fully polished. The papers carry the names of the authors and should be cited accordingly. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the International Bank for Reconstruction and Development/World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent.

Produced by the Research Support Team

A Vulnerability Approach to the Definition of the Middle Class[1]

Luis F. Lopez-Calva
The World Bank

Eduardo Ortiz-Juarez[2]
Economist, RBLAC-UNDP and Consultant, The World Bank

JEL: D31, I30, D63
Key Words: Middle class; Income distribution; Poverty; Vulnerability; Panel data.
Sector Board: POV

[1] We would like to thank Paloma Anos-Casero, Nancy Birdsall, François Bourguignon, Francisco Ferreira, Peter Lanjouw, Santiago Levy, Maria Ana Lugo, Nora Lustig, Julian Messina, Jamele Rigolini, Carlos Silva-Jauregui and the team of the mobility flagship in OCE-LAC at The World Bank, as well as participants at the 15th LACEA Meeting held in Medellin, for helpful comments. Megan Rounseville provided invaluable research and editing support. Any errors or opinions are the authors' and do not necessarily represent the views of the World Bank or those of the United Nations Development Programme.
[2] Luis Felipe Lopez-Calva (lflopezcalva@worldbank.org) is the Lead Economist for the Poverty, Equity and Gender Unit, Latin America and the Caribbean at the World Bank; and Eduardo Ortiz-Juarez (eduardo.ortiz@undp.org) is an Economic Analyst at the UNDP's Regional Bureau for Latin America and the Caribbean (UNDP-RBLAC) and Consultant at the World Bank.

1. Introduction

The concept of social class and specifically middle class, has been widely discussed in sociology and other social sciences, but mostly ignored in modern economics. In practice, the middle class has been defined in terms of income, consumption patterns, occupational status, or even by using self-identification. Regardless of which definition is used, the measurement of middle class is dependent on a particular period and place and it is determined by several factors, such as history, culture and the development stage of a society. Analysis of the middle class emerges as a central issue given the strong influence this social group has on society, politics and the economy. Much literature (Easterly 2001) suggests that the middle class helps to produce economic benefits and foster economic development, through its emphasis on human capital investment, consumption and savings, which, in turn, incentivizes a virtuous circle contributing to further expansion of this social group. Other authors (Birdsall 2010) suggest that the middle class constitutes the backbone of democracy ensuring social and political stability by fostering social cohesion and mitigating tensions between the poor and the rich. As political scientists suggest, a strong and stable middle class is usually accompanied by a more "stable democracy".

Over the past decade several countries in Latin America have accomplished important reductions in poverty and inequality and as a result policymakers are now confronting new challenges, and seek to understand problems they face from a political economy perspective. The simultaneous emergence of improved data quality and availability has sparked a surge in literature addressing measurement and dynamics of the middle class. However, most writing on the subject lacks clarity on the definition of what it means to be *middle class*. In the sociological literature there is a long tradition of class dynamics analysis based mainly on occupational structure following Goldthorpe's categories (Goldthrope 1987). However, in the economic literature, the analysis has focused mainly on relative definitions, addressing a *stratum* of the income distribution rather than an analysis of *class*. Existing relative definitions compare different middle classes from place to place because income distributions differ across countries. An absolute approach becomes more advantageous because it identifies middle class as those households with income or consumption in a specific and comparable range. For instance, Banerjee and Duflo (2008) (B&D from here onwards) and Ravallion (2010) have suggested the use of absolute income thresholds to define lower, middle, and upper classes. However, while these absolute measures enable comparison across countries their definitions have resulted in descriptive statistics of income groups because the thresholds are defined arbitrarily.

The response to the question of why we care about the measurement, the analysis and the empirical contrast of different definitions of the middle class can found somewhere else (Birdsall et al. 2011; Hertova et al. 2011; Cruces et al. 2010). In line with Amartya Sen's (1983) statement: "poverty is absolute in the realm of capabilities but relative in the realm of income", we propose a framework in which middle class is absolute in terms of the *functionings* that define it but relative in terms of the means through which those functionings can be achieved. In this paper we argue that vulnerability to poverty is the absolute functioning that defines the

middle class. We set the lower threshold of the middle class at an absolute level: a 10 percent probability of falling into poverty. Using a regression-based approach we exploit panel data to determine the amount of comparable income associated with that probability level –using income as the relative measure of vulnerability to poverty. Based on our findings from applying this methodology to three countries, we set an absolute lower bound for the middle class of 10 dollars PPP. This creates an absolute lower threshold that can be used to measure the middle class across countries over time. We, then, apply the 10-dollar PPP absolute lower threshold to cross-sectional surveys (household data from 1992 through 2008-09) to measure the size of the middle class in Chile, Mexico and Peru and analyze its evolution over the last two decades. Lastly we contrast this new absolute definition with existing absolute definitions of the middle class presented by B&D and Ravallion.

Evidence from the application of the vulnerable-to-poverty approach shows that both the proportion of middle class households and the income share appropriated by this group have significantly increased in all three countries during the period under study –in-tandem with economic growth in these countries as we should expect. We also find that the proposed lower threshold of middle class is at or above the median of the income distribution in the countries analyzed. We show that alternative absolute definitions (B&D and Ravallion) lump people who are still vulnerable to poverty into their definition of middle class, which has resulted in counter-intuitive trends in the size of the middle class –the middle class expands during economic downturns and shrinks in times of growth– a problem the vulnerability-to-poverty measure corrects.

In summary, the vulnerable-to-poverty approach defines the middle class in absolute terms as those households with income or consumption in a specific and comparable range, and it makes two important contributions to the measurement of the middle class. First, methodologically, it ensures that no lower class or poor households are being identified as middle class. Second, conceptually, it establishes a well-defined conceptual framework for the analysis of the middle class over time.

This paper is structured as follows. Section 2 provides a brief literature review on the relationship between vulnerability and middle class. Section 3 presents an overview of existing measures of middle class and their shortcomings. Section 4 describes the data used in this paper. Section 5 describes the empirical strategy to estimate an absolute-standard for middle class analysis. Section 6 presents the results and contrasts it with other existing absolute definitions. Finally, section 7 concludes.

2. Linking the notion of vulnerability to the definition of middle class

The concepts of class and middle class have been widely discussed in sociology and other social sciences, but mostly ignored in modern economics. The conceptual roots of class can be found in the work by Carl Marx and Max Weber. The Marxian approach defined class in terms of common structural positions within the organization of production, where class stratification is based on the concept of exploitation and property relations (Wright, 1979). The Weberian view, on the other hand, conceived class as individuals with common economic "life

chances" which determine their opportunities for income in the market, and identified the middle class as those owning skills and education. This approach has triggered an important interest in analyzing the link between class and economic vulnerability. In a recent study, Goldthorpe and McKnight (2004) show that class positions affect the risk and opportunities individuals face. Their analysis focuses on three classes of workers and their contracts (non-skilled workers with simple contracts, professional workers and managers with comprehensive and stable contracts, and intermediate workers with "mixed" forms of contracts), as well as on three dimensions: economic security, economic stability and economic prospects. The main results from the empirical work show that, in terms of economic security, the type of the contracts held by non-skilled workers has a direct impact on the high risks of unemployment and job loss these workers face, relative to other classes. In this sense the "commodifying" of labor and associated higher probability of unemployment and less security in terms of health and pensions, makes class a central element of economic vulnerability.

On the issue of vulnerability, there is vast literature discussing and measuring the concept. In economics this is mainly done from the perspective of poverty traps (Ligon and Schechter 2004; Hoddinott and Quisumbing 2003) and poverty dynamics (Dercon 2006, 2001; Elbers and Gunning 2003). In Dercon's work, the author develops a framework for analyzing poverty and its links with vulnerability focusing on households and distinguishing three levels: assets, incomes generated from these assets, and capabilities. The main argument is that transitions from one level to the next (i.e. the transformation of assets into incomes and in turn their transformation into well-being) involve different risks (environmental and health factors, market- and public-related risks), for which both household strategies and public and private actions try to cope.

The relevance in analyzing vulnerability is related to the growing interest in including the concept into the design and implementation of social development strategies. Under the normative assumption that vulnerability causes insecurity and a reduction of welfare, it should be a central component of public actions to ensure a minimum level of protection. In an empirical work using panel data for rural India, Gaiha and Imai (2008) employ a decomposition of both ex ante and ex post measures of vulnerability to aggregate and idiosyncratic risks. An important result is that small farmers are vulnerable to idiosyncratic risks, which cause them to reduce their levels of consumption. Furthermore, vulnerability must be viewed as a broadening area for public intervention (that goes beyond static poverty status) given the presence of inter-temporal risks such as changes in income levels and sources, climate-related events, and conflicts, particularly important as markets for credit and insurance are incomplete and imperfect.

While there is widespread consensus on the inclusion of vulnerability as a component of development policies, the question remains on how to measure it and how to quantify its impacts on welfare.[3] Recently, Cafiero and Vakis (2006) have suggested an approach based on an "augmented" poverty line, which in addition to including a bundle of consumption goods

[3] The literature has proposed a number of measures designed to capture the welfare consequences of vulnerability for poor households. See for example the discussion by Ligon and Schechter (2004) and Dercon (2002).

and services it also incorporates a basic "basket of insurance" against "unacceptable risks". The main argument is that by embedding risk in the measurement of poverty, such a measure becomes a wider concept of vulnerability that captures the welfare consequences of exposure to risk, not only that of having been subject to shocks.

Longitudinal data for Chile, Mexico and Peru –described in detail in section 4— allow us to observe the incidence of events that could potentially affect welfare. In Mexico, table 1 shows the incidence of several shocks between 2002 and 2005: death or illness of a household member; economic shock defined as business bankruptcy or unemployment; and loss of dwelling, crops and livestock due to climate-related events. While a higher incidence of loss of crops and livestock is evident for poorest households (1st quintile) –probably due to these activities are mostly rural and where poor population could be more exposed to shocks— an important result is that the occurrence of deaths (8.8 percent), health shocks for which the individual was hospitalized (12 percent), economic shocks (6.6 percent) and, in general, any kind of shock (25 percent), is a somewhat stable through the income distribution suggesting thus that the entire population is prone to negative shocks.

Table 1: Incidence of shocks by income quintiles; Mexico 2002-2005

Percentage of households

Income quintiles	Any shock	Death	Health shock	Economic shock	Loss of dwelling	Loss of crops	Loss of livestock	Dwelling, crops and livestock
1	29.0	10.8	11.5	6.1	1.6	5.7	2.3	8.0
2	24.9	8.8	12.0	6.5	1.3	3.7	1.0	5.4
3	25.8	8.7	12.2	7.4	0.4	1.9	0.5	2.5
4	24.2	8.2	12.1	7.1	0.5	1.2	0.7	2.2
5	21.5	7.4	11.0	6.0	0.5	1.5	1.5	3.1
Total	**25.1**	**8.8**	**11.8**	**6.6**	**0.9**	**2.8**	**1.2**	**4.2**

Source: Author's calculations based on data from MxFLS panel database.

In the cases of Chile and Peru the data shows only the occurrence of health shocks, so that a cross-country comparison is only possible for those shocks requiring hospitalization. While the occurrence of health shocks is stable for all five quintiles in Mexico, in Chile and Peru it shows a growing pattern through the income distribution. Although this pattern suggests a higher use of hospitalization for upper quintiles according to the capacity to pay, these results evidence a generalized exposure to risks in all three countries (see table 2).

Given that insurance markets are actually imperfect or inexistent, households insure against risks through less efficient strategies (e.g. investing resources in activities with low returns, but which assure a minimum level of consumption), making the efficient cost of insuring against risk both variable and problematic to realistically define. We argue that it is possible to find the income level associated with a set of assets and socioeconomic characteristics that would allow the households to be less vulnerable to fall into poverty due to idiosyncratic and asymmetric shocks, and interpret it as an "augmented" poverty line in the spirit of Cafiero and Vakis. From that perspective, and based on the notion discussed in Goldthorpe and McKnight, we explore the link between income and vulnerability to poverty to define the lower threshold of

middle class for three Latin American countries. The conceptual basis follows the vulnerability to poverty approach, by which, the middle class is defined the level of income that allows individuals to protect themselves from falling into poverty over time.

Table 2: Incidence of health shocks; 5-year period: Chile 2001-06, Mexico 2000-05, and Peru 2001-06[/a]

Percentage of households

Income quintiles	Chile	Mexico	Peru	
			At least one member	Head
1	11.1	11.5	30.7	7.2
2	11.6	12.0	39.8	10.0
3	14.2	12.2	44.3	12.5
4	15.1	12.1	46.3	15.0
5	18.2	11.0	43.6	17.1
Total	**14.0**	**11.8**	**40.9**	**12.4**

Source: Author's calculations based on data from CASEN, MxFLS and ENAHO panel databases.

[/a] The differences in the incidence for Chile and Mexico with respect to Peru is probably due, among other factors, to the way the surveys collects the information (first three columns). In Chile and Mexico the surveys in 2006 and 2005, respectively, asks for the occurrence of health shocks during the previous 5 years for which household members required hospitalization. The 5-year period probably underestimates the incidence as individuals are more likely to forget the occurrence of shocks. In Peru, however, the surveys in 2002, 2003, 2004, 2005 and 2006 ask for the occurrence of these shocks during the last 12 months, making it more likely to capture all the events.

3. Measurement of the middle class

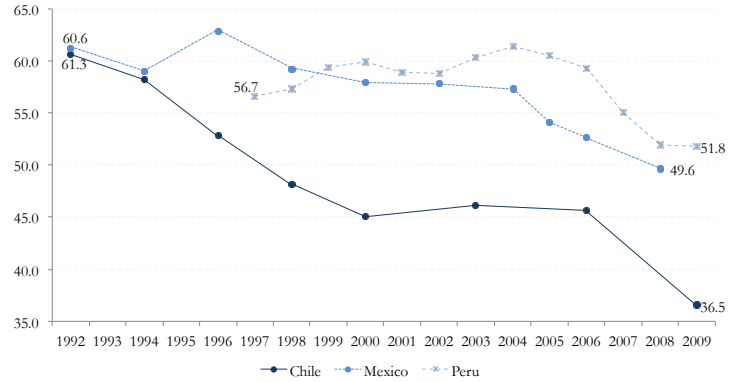
Existing income or consumption thresholds used to measure and analyze trends for the middle class have been defined using ad-hoc criteria. Milanovic and Yitzhaki (2002) have defined the middle class as those individuals living with a per capita income on $12-50 a day at 2005 PPP (roughly the mean per capita incomes of Brazil and Italy, respectively), while Kharas and Gertz (2010) have used a range of $10 and $100 daily expenditure per person, obtained after adjusting survey distributions to national accounts. These thresholds were defined by excluding those individuals who would be considered rich in Portugal and poor in Luxembourg (the poorest and richest among the industrialized countries, respectively).

Moving to a developing world perspective, Birdsall (2010) defined the middle class to include those with income between $10 day and at or below the 95th percentile of the income distribution. She argues that $10 a day is a high income compared to the global absolute poverty line of $1.25 a day and, although it is still low by OECD standards, it is close to setting a standard for a minimum income threshold for "economic security". In an influential work, B&D have defined middle class as those living with a per capita expenditure on $2-10 a day at PPP, while Ravallion proposed thresholds ranging between $2 a day (the median poverty line in 70 developing countries), and $13 a day (the poverty line in the United States). Both Ravallion and B&D's measures have triggered an important debate on internationally comparable measurement of the middle class. The use of these ad-hoc criteria, when analyzing

trends of the middle class over time, produces outcomes that don't align with our general conceptual definition of the middle class. Applying the thresholds defined by Ravallion and B&D to Chile, Mexico and Peru we observe that the middle class has decreased over the whole period of time analyzed in this paper. Conversely, the middle class has increased in times of economic downturn moving in parallel with poverty levels.

Figure 1: People living on $2-10 a day in Chile, Mexico and Peru (B&D's measure of middle class)

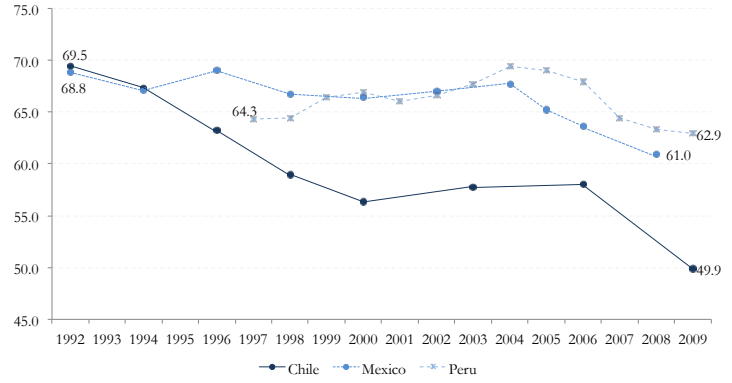
Percentage of households



Source: Author's calculations based on data from CASEN, ENIGH and ENAHO cross-sectional databases.

Figure 2: People living on $2-13 a day in Chile, Mexico and Peru (Ravallion's measure of middle class)

Percentage of households



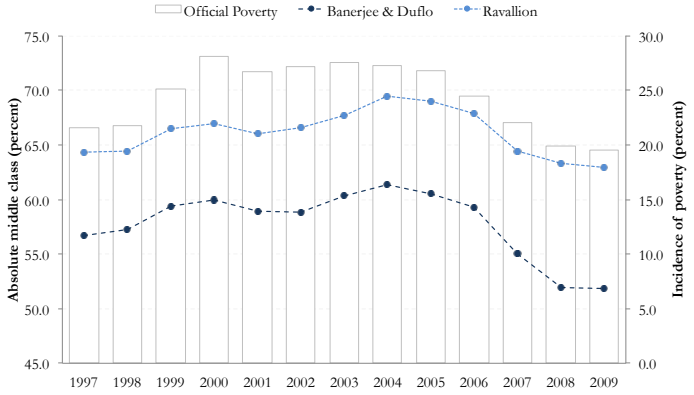
Source: Author's calculations based on data from CASEN, ENIGH and ENAHO cross-sectional databases.

As figures 1 and 2 show, the size of the middle class increased in Mexico during the *tequila crisis* of 1995 because the incidence of poverty also increased from 23.2 percent in 1994 to 38 percent in 1996. In Peru, the economy suffered a triple blow during 1998-99 with President Fujimori's political collapse, the financial turmoil in BRIC's countries, and a serious damage caused by *El Niño* phenomenon. As a result, the incidence of poverty increased from 21.8 percent in 1998 to 25.1 percent in 1999 and to 28.1 percent in 2000; surprisingly the middle class also increased. In the following years, in spite of Peru's economic performance being among the best in Latin America since 2002, the unemployment rates were still high and

poverty slightly increased after a slow drop in 2001; again, the size of the middle class increased. Overall, these figures suggest that the evolution of the absolute middle class has followed the same pattern than the official poverty incidence (see figure 3), which is both counter intuitive and contradictory to the definition of middle class presented here. As we argue, the lack of a conceptual framework to support a definition of thresholds results in descriptive statistics rather than meaningful analysis.

Figure 3: People living on $2-13 a day; on $2-10 a day; and incidence of poverty in Peru[/a]

Percentage of households



Source: Author's calculations based on data from ENAHO cross-sectional databases.
[/a] The left axis draws the size of the middle class according to B&D and Ravallion, while the incidence of poverty is drawn in the right axis.

Atkinson and Brandolini (2011) have a consistent view in the sense that they suggest that income distribution and income measures alone can no longer represent accurate definitions of the middle class. They argue for the inclusion of wealth holdings and position in the labor market into income definitions of class. The latter has been the longest tradition in sociological work.

4. The data

The vulnerability-to-poverty approach to defining the middle class is constructed in three stages. The first stage identifies actual characteristics that are associated with movement in or out of poverty. Then, the second and third stages construct probabilities of falling into poverty, and an income level associated to those, respectively. Based on such models, the methodology results in an estimate of the value of the lower income threshold, in PPP terms. In these stages we exploit longitudinal data for Chile (2001-2006), Mexico (2002-2005), and Peru (2002-2006).

In the case of Chile, we use the panel version of the *Socioeconomic Characterization Survey* (CASEN Panel) carried out by the Ministry of Planning (MIDEPLAN). The first wave was undertaken in 1996 using a random sub-sample of 5,210 households from the cross-sectional CASEN covering four regions of the country, which represent 60 percent of the national

population. Approximately 15,038 household members in the baseline were re-interviewed in the wave of 2001, and 12,100 of these were surveyed in 2006. These figures imply an approximately 20 percent attrition rate between 2001 and 2006, excluding all new households that branched out from the 2001 original ones. For Mexico, the data is taken from the *Mexican Family Life Survey* (MxFLS) for the first (2002) and second (2005) rounds. The survey is representative at the national, regional, urban and rural levels. The first wave includes approximately 8,440 households distributed in 150 localities, and the second wave includes 7,572 of the original households for an attrition rate of only 10 percent. As before, we only consider those households interviewed in both 2002 and 2005, excluding all new households that branched out.

Finally, for Peru we use the panel of the *National Household Survey* (ENAHO Panel) collected by the Institute of Statistics (INEI), providing a five-wave unbalanced panel for the period 2002-2006. The size of the sample (6,257 in 2002; 4,217 in 2003; 6,490 in 2004; 6,778 in 2005; and 6,593 in 2006) is representative at the national, urban, rural, and regional levels (Coast, *Sierra* and *Selva* regions). The unbalanced nature of the panel results in gaps in the observations, since there is no information for all five waves for every household. Therefore, in this paper we use the information of those households interviewed in both 2002 and 2006 (5,092 households), which imply an attrition rate of approximately 20 percent similar to that found in the case of Chile, and comparable to longitudinal surveys in other regions of the world.

Then, the definition of the middle class with the vulnerability-to-poverty approach is applied to obtain the lower PPP income threshold to cross-sectional data in order to analyze the evolution of the middle class over a longer period of time. For Chile (1992-2009), the data is taken from the cross-sectional CASEN, carried out every two years since 1985 by MIDEPLAN. This data source is a nationally and regionally representative household survey used to measure the incidence of poverty and the level of inequality in the country. For Mexico (1992-2008), we use the *National Household Consumption and Expenditure Survey* (ENIGH), undertaken by the Institute of Statistics and Geography (INEGI). The ENIGH is a nationally representative survey, covering urban and rural areas, and it contains detailed information on income and different categories of expenditures. This survey is available for 1984 and 1989, and for every two years since 1992. Finally, in the case of Peru (1997-2009) we use data from the cross-sectional ENAHO, collected on a quarterly basis by INEI since 1997. The fourth wave of the survey is nationally representative, and it is used for poverty measurement as it collects information on income and expenditure. All three surveys collect information of dwelling characteristics, and indicators related to age, sex, employment, and education of the household members.

5. A vulnerability approach to middle class analysis

We propose an empirical framework to define the middle class in terms of households' vulnerability to poverty following a regression-based approach to estimate an income lower threshold associated to a low probability of falling into poverty, which will be used for cross-

country comparisons.[4] Methodologically, we proceed in three stages. First, we construct poverty transition matrices from panel data using the official moderate poverty lines in Chile, Mexico and Peru, which range between $4 and $5 a day at 2005 PPP.[5] These matrices shows a cross-classification of poverty status in the initial time point in the rows, and the poverty status in the final time point in the columns (see table 3), so that the proportional distributions across rows provide an assessment of the transition to poverty and the extent of poverty stability.

Table 3: Poverty transition matrices

Percentage of households

a) Chile: 2001-2006

		Final Period →		
Initial Period ↓		*Non-poor*	*Poor*	*Total*
	Non-poor	92.0	8.0	100
	Poor	66.2	33.8	100

b) Mexico: 2002-2005

		Final Period →		
Initial Period ↓		*Non-poor*	*Poor*	*Total*
	Non-poor	81.7	18.3	100
	Poor	50.9	49.1	100

c) Peru: 2002-2006

		Final Period →		
Initial Period ↓		*Non-poor*	*Poor*	*Total*
	Non-poor	82.1	17.9	100
	Poor	39.2	60.8	100

Source: Author's calculations based on data from CASEN, MxFLS and ENAHO panel databases.

This analysis allows us to classify households into four categories: 1) a household is classified as *never poor* if it has never fallen under the poverty line in the two periods of each panel; 2) conversely, it is classified as *always poor* if it has been poor in both waves of the panel; 3) households are also classified as *out of poverty* if they were poor in the initial time period, but exit out of poverty in the final period; 4) finally, households that were non-poor in the initial period but fell into poverty in the final year are classified as *entering poverty*.

The idea behind calculating the actual transitions is that rather than looking at actual income we interpret it as a summary metric that gauges the level of household's asset ownership, and construct measures of permanent income by regressing income on assets and several socioeconomic characteristics. We argue that households' ability to cope with shocks is as much dependent on their actual income, than on wealth and their ability to cope with risks. Therefore, while because of data constraints the middle class threshold must ultimately be expressed in terms of income, looking at transitions using permanent income appears to be more appropriate.

[4] A variation of the proposed methodology using cross-sectional data can be found in Hertova *et al.* (2011).

[5] It refers to moderate poverty lines (*pobre no indigente* in the cases of Chile and Peru, and *pobreza de capacidades* in the case of Mexico). All the estimations are strongly consistent if we use the international poverty line of $4 a day instead of the official ones.

In a second stage, we move to a logistic model in order to analyze the correlates of the probability of falling into poverty over the analyzed period. The estimated probability of being poor in the next period p_{it} for a household i is given by:

$$p_{it} = E\left(poor_{i,t+1}|\mathbf{X}_{it}\right) = F(\mathbf{X}_{it} \cdot \boldsymbol{\beta}_{it}) \quad (1)$$

where $poor_{i,t+1}$ is the dependent variable taking the value of 1 if households are identified as *always poor* in both periods or *entering poverty* in the final period (accounting for 14, 26.2, and 39.2 percent of the sample in Chile, Mexico and Peru, respectively) and 0 otherwise; $\boldsymbol{\beta}_{it}$ is a vector of model parameters, and $\mathbf{X}_{it}$ is a vector of observable characteristics including demographic indicators, labor market resources, and shocks affecting the household. For the initial time point, the demographic characteristics include rural residence, as well as age, sex and marital status of the household head. The last variable identifies if the head is married, cohabiting, or single, including for this category if she/he is widowed, separated or never married.

For market resources we include the head's education level as a proxy for human capital. In Chile and Peru, we use an ordinal variable with seven categories: no formal education; incomplete and complete primary education; incomplete and complete secondary education; and incomplete and complete tertiary education. In Mexico, we distinguish six categories: no formal education; incomplete and complete primary education; complete lower-secondary; complete upper-secondary; and some year of tertiary education or more. In addition, we include the occupational status of the head in the model. In Chile, we use a collapsed six-class version of the EGP class classification (Erikson et al. 1979): professional and managers, clerical workers, self-employed, skilled manual workers, non-skilled manual workers, and agricultural workers. In Mexico, we use two additional categories: workers engaged in commerce and sales, and army and police officers. The classification used in Peru considers those workers in agriculture, energy, manufacturing, construction, commerce, communications, and clerical activities. The model also accounts for changes in the number of household members engaged in paid employment within the period considered, as well as in the household size.

Finally, a self-reported indicator for economic shocks affecting households is included. In Chile and Peru we include the incidence of health shocks requiring hospitalization during the period under study, while in Mexico the variable we use includes death, illness or accident of any household member, the economic failure of a household member (unemployment and bankruptcy of familiar business), and the loss of housing, business, crop and livestock due to climate-related events.

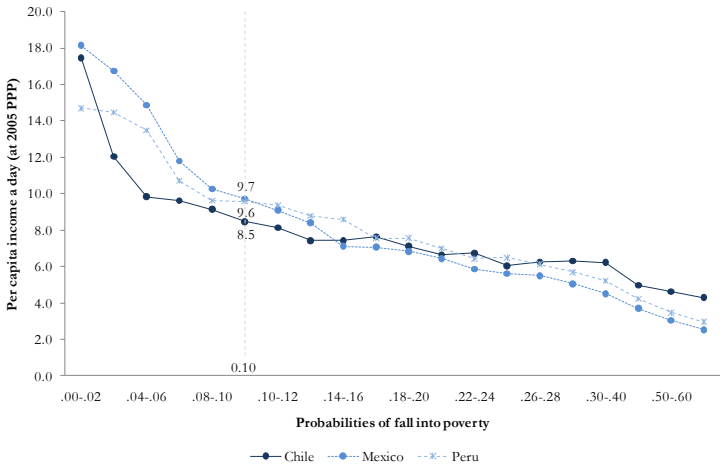
In the third stage, we calculate the average of the independent variables for an array of estimated probabilities of falling into poverty. Then, we use the same independent variables in equation (1) to estimate an income equation expressed as:

$$lnY_{it} = \alpha + \mathbf{X}_{it} \cdot \boldsymbol{\beta}_{it} + \varepsilon_i \quad (2)$$

where lnY_{it} is the household per capita income in logarithmic scale at the initial time point. The resulting coefficients from equation (2) and the average of the independent variables are used to solve the income equation, and therefore to obtain the amount of income associated to each probability.

The resulting incomes look quite similar for all three countries; non-poor individuals with a ten percent probability of falling into poverty have income levels at $8.5 a day in Chile, $9.7 in Mexico, and $9.6 a day in Peru, all of them measured using the 2005 PPP (see figure 4).[6]

Figure 4: Daily income by probabilities of falling into poverty; Chile, Mexico and Peru
Household per capita income at 2005 PPP



Source: Author's calculations based on data from CASEN, MxFLS and ENAHO panel databases.

These figures result from the complete specifications of the models shown in annex (table A4); however, they are robust to different specifications (see table 4). As the middle class, ideally, should consist of those households facing a very low risk of falling into poverty over time, we define the income associated to that probability as the lower threshold that depicts the lower bound of the middle class.[7] While there is a strong association between income and vulnerability, it remains nevertheless difficult to anchor a threshold solely to vulnerability, since the curves do not suggest structural behavioral changes. In the case of the upper threshold we define it as $50 in PPP terms, being it an income amount that lies in the upper tail of the income distribution of all three countries.

[6] Estimated as the average of the independent variables for an estimated probability range between 0.09 and 0.11. Tables A1-A3 in the annex show the coefficients and standard errors from equations (1) and (2) for Chile, Mexico, and Peru, respectively, while Figure A1 shows the correlation between the estimated probabilities from equation (1) and the fitted values from equation (2).

[7] In order to not classify as middle class a number of lower class households if any of these incomes is used, we establish for strict comparability purposes a more demanding criterion of $10 a day. Of course, nothing comes without a trade-off. The use of the proposed threshold of $10 implies, conversely, that a number of middle class households in each country will be considered as lower class. For instance, if middle class in Mexico includes those households with per capita income at or above $9.7 a day, hence a 1.3 percent of these households will be lower class under the criterion of $10 (in Chile 8.2 percent of households are between $8.5 and $10, while in Peru 1.3 percent are between $9.6 and $10). However, we consider more reasonable to include a lower proportion of middle class households as part of the lower class, instead of the opposite situation.

The definition of the upper threshold is not the focus of this paper and it is less relevant than the lower threshold because of three main factors. First, because moving the upper threshold up (down) the income distribution includes (excludes) a small percentage of individuals. As shown by Birdsall et al. (2011), varying the upper threshold from $50 to $100 a day would move the percentile of the LAC's elite from the top 2.2 percent to the top 0.5 percent. In contrast, moving the lower threshold would move dramatically the percentage of excluded/included population. Second, an income threshold above $50 would also limit the representativeness of the upper class in some countries. Finally, and related to the second factor, an upper threshold of $50 seems reasonable as it has been evidenced that household surveys do not collect realistic information of the richest population of a country (Alvaredo and Piketty, 2010), so that it is likely that a number of households in the surveys' top percentiles would be considered, actually, as part of the middle class. Overall, we propose that the middle class is thus defined as those households living with a per capita income on $10-50 a day at 2005 PPP.

Table 4: Robustness to different models specifications

Daily per capita income for a ten percent probability of falling into poverty

Model Specification	Chile	Mexico	Peru
Complete specification (see annex)	**$ 8.5**	**$ 9.7**	**$ 9.6**
Excluding shocks	$ 8.5	$ 9.9	$ 9.5
Excluding shocks and variables measuring changes	$ 8.8	$ 11.1	$ 11.0
Excluding variables measuring changes	$ 8.6	$ 10.9	$ 10.9
Excluding locational effects, shocks and changes	$ 9.6	$ 11.3	$ 11.3
Excluding locational effects, marital status, shocks and changes	$ 9.0	$ 11.4	$ 11.3
Only education, sex, age and occupational status of heads	$ 9.8	$ 11.5	$ 11.8
Only education, sex, and age of heads	$ 9.6	$ 12.0	$ 11.7

Source: Author's calculations based on data from CASEN, MxFLS and ENAHO panel databases.

In order to observe how the socioeconomic characteristics of the middle class compare relative to other groups, we estimate a two-group mean-comparison *t*-test to observe if there is a significant difference between the characteristics of the middle class and those of the lower class. For that purpose, we first divide the lower class in two groups: i) *lower* lower-class, or poor class, defined as those households with per capita income below the official poverty lines as measured in each country, and ii) *upper* lower-class, or vulnerable class, defined as those households with per capita income between the poverty line and $10 a day. Second, we compare each of these groups with the middle class households living on $10-50 a day. In an attempt to define a standard profile of the middle class in all three countries, table A5 in the annex shows that the middle class is characterized as having higher levels of schooling, especially tertiary education; living in primarily urban areas, and its heads perform clerical or professional activities.

Differences between the middle class and other groups are large and statistically significant. The three variables that show the largest differences among the group means are percentage rate of households with dirt floors, the head of the household occupation is a farmer and rural residence are (see table 5). In each case the poor classes have the highest prevalence rates

followed by the vulnerable classes and lastly by the middle classes. The considerable and statistically different differences between the vulnerable class and the middle class emphasize the importance of making the distinction of vulnerable households from the middle class.

Table 5: Select statistically significant mean comparisons[/a]

Percentage of households

	Poor	Vulnerable	Middle
		Chile	
Dirt floor	10.7	6.4	3.2
Head is farmer	40.3	23.4	12.9
Rural	33.8	22.7	13.5
		Mexico	
Dirt floor	24.8	10	3.5
Head is farmer	48.4	19.5	10.3
Rural	76.8	48.5	31.5
		Peru	
Dirt floor	69.4	34.2	17.4
Head is farmer	71.6	31.5	16.6
Rural	77.5	44.8	26.8

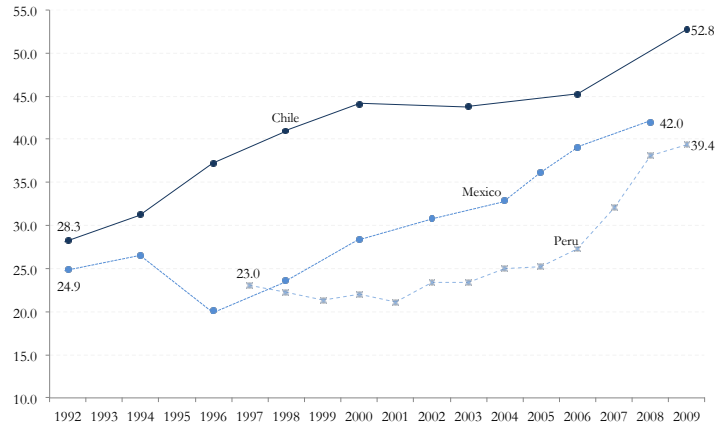
Source: Author's calculations based on data from CASEN, MxFLS and ENAHO panel databases.
[/a] Full table with t-tests in annex; table A5.

6. The size of the middle class over the last two decades

Applying our absolute definition to cross-sectional surveys, the results show that the percentage of middle class households has increased significantly from 28.3 to 52.8 percent in Chile, from 24.9 to 42 percent in Mexico, and from 23 to 39.4 percent in Peru over the analyzed period (see figure 5). It is important to notice that the growth of the middle class has been sustained in Chile over the complete time span, while in Mexico and Peru this occurred only after the economic downturns of 1994/95 in Mexico, and the downturns of 1998/99 and 2001 in Peru.

Figure 5: Size of the middle class in Chile, Mexico and Peru

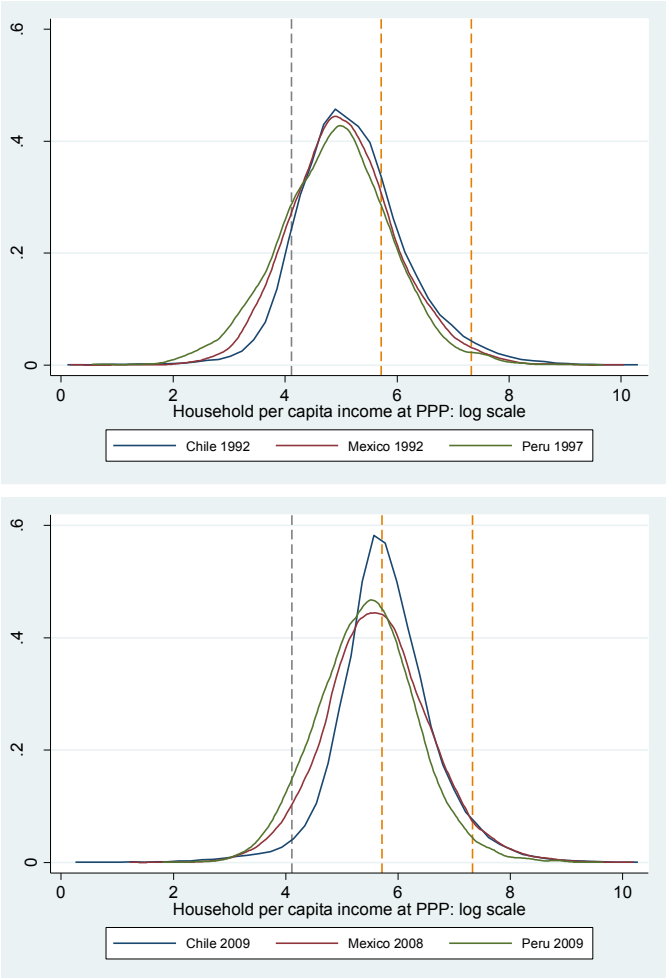
Percentage of households



Source: Author's calculations based on data from CASEN, ENIGH and ENAHO cross-sectional databases.

The middle class has not only increased in size, but it has also increased the income share appropriated by this group. In Chile, this share increased 13.1 percentage points from 42.6 percent in 1992 to 55.9 percent in 2009; in Mexico it increased 7.8 percentage points from 45.2 to 52.9 percent over 1992-2008, and in Peru it expanded by 13.1 points from 42.7 to 55.7 percent over 1997-2009. These results become clear when looking at the income distribution. Figure 6 shows that the bulge in the middle increased over the analyzed period making the bell taller and confirming the increase of the middle class, partly due to improvements in income distribution (motivated by an expansion and enhancement of social policies) and a reasonable economic performance, which is evident by the movement of the bell to the right.

Figure 6: Kernel distributions of income in Chile, Mexico and Peru



Source: Author's calculations based on data from CASEN, ENIGH and ENAHO cross-sectional databases.
/a The gray line is a lower threshold of $2 a day proposed by B&D and Ravallion; while the yellow lines are the thresholds proposed in this paper.

Overall, the evidence found suggests a higher number of middle class households facing a low probability of falling into poverty in the late-2000s than in the 1990s, which is undoubtedly

very important in light of the reasons outlined in a large literature arguing that size and composition of the middle class is critical for strengthening and stabilizing the democratic system and its institutions, for a better economy's performance, and for social cohesion. Moreover, while our definition follows a vulnerability-to-poverty approach and empirically establishes a new lower threshold for middle class identification, other existing definitions have tried to answer the question of what income level to choose for middle class measurement, as discussed above. We contrast these definitions with our proposed range and found that the middle class has increased in Chile, Mexico and Peru during the period under study, contrary to the results presented from the application of the Ravallion and B&D methods (shown previously: see figures 1-2).

The divergences are rather evident: the absolute lower threshold defined by Ravallion and B&D ($2 a day) is a very low standard for middle class comparisons across the upper-middle income countries analyzed in this paper. Actually, the official poverty lines in all three countries are more than twice the $2 threshold, suggesting thus the inclusion of poor households as part of the middle class (see Kernel distributions in annex; figure A2). The latter implies that the decrease in the size of the middle class, according to B&D and Ravallion definitions, was motivated by the fact that a proportion of households climb out of poverty during the analyzed period: in Chile from 32.9 to 15.1 percent over 1992-2009; in Mexico from 23.1 to 20.1 percent over 1992-2008; and in Peru from 21.6 to 19.6 percent over 1997-2009. The vulnerability-to-poverty measure of middle class proposed in this paper corrects these trends.

7. Concluding remarks

This paper proposes an empirical methodology to analyze the middle class based on the notion of vulnerability to poverty. Following a regression-based approach this paper exploits panel data to estimate the amount of comparable income –$10 dollars a day, associated with a low (0.10) probability of falling into poverty— which depicts the beginning of the middle class (lower threshold). In the case of the upper threshold it is established at $50 dollars a day. By doing so, this paper proposes an absolute-standard definition of the middle class as those households living with a per capita income of $10-50 a day at 2005 PPP. These thresholds are then used in cross-sectional surveys to measure the size of the middle class in Chile over 1992-2009, in Mexico over 1992-2008, and in Peru over 1997-2009.

Following this definition, this paper overcomes the two main shortcomings that emerge when comparing middle classes across countries using existing measures. First, relative measures face the problem of different income or consumption distributions from one country to another, and thus they compare different middle classes. Second, the use of very low absolute thresholds in upper-middle income countries, such as Chile, Mexico and Peru, includes a high proportion of poor households within the middle class definition. This sheds light on the descriptive analysis that emerges from the use of monetary thresholds without a conceptual framework. Overall, this paper makes two important contributions to the measurement of the middle class. First, methodologically, it ensures that no lower class or poor

households are being identified as middle class. Second, conceptually, it establishes a well-defined conceptual framework for the analysis of the middle class over time. The evidence suggests a higher number of middle class households facing a low probability of falling into poverty in the late-2000s than in the 1990s, which is undoubtedly very important in light of the reasons outlined in a large literature arguing that size and composition of the middle class is critical for strengthening and stabilizing the democratic system and its institutions, for better economic performance, and for social cohesion.

References

Alvaredo, F. & T. Piketty (2010). "The Dynamics of Income Concentration in Developed and Developing Countries: A View from the Top", in Lopez-Calva, L.F. & N. Lustig (eds.), *Declining Inequality in Latin America: A Decade of Progress?* Baltimore: Brookings Institution Press.

Atkinson, A.B. & A. Brandolini (2011). "On the Identification of the 'Middle Class'", Paper presented at the fourth meeting of the Society for the Study of Economic Inequality (ECINEQ). Catania, July.

Banerjee, A. & E. Duflo (2008). "What Is Middle Class about the Middle Classes around the World?" *Journal of Economics Perspectives*, 22(2): 3-28.

Birdsall, N. (2010). "The (Indispensable) Middle Class in Developing Countries; or, the Rich and the Rest, Not the Poor and the Rest", Working Paper 207. Center for Global Development, Washington, D.C.

Birdsall, N., F. Ferreira, L.F. Lopez-Calva & J. Rigolini (2011). "The Middle Class in Developing Countries: Concept, Measurement and Recent Trends", Mimeo. The World Bank, Washington, D.C.

Cafiero, C. & R. Vakis (2006). "Risk and Vulnerability Considerations in Poverty Analysis: Recent Advances and Future Directions", Social Protection Discussion Paper 0610. The World Bank, Washington D.C.

Cruces, G., L.F. Lopez-Calva & D. Battiston (2010). "Down and Out or Up and In? In Search of Latin America's Elusive Middle Class", Inclusive Development Series No. 3. RBLAC-UNDP, New York.

Dercon, S. (2006). "Vulnerability: A Micro Perspective", QEH Working Paper Series. University of Oxford, Oxford.

Dercon, S. (2002). "Income Risk, Coping Strategies and Safety Nets", Discussion Paper No. 2002/22. UNU-WIDER, Helsinki.

Dercon, S. (2001). "Assessing Vulnerability to Poverty", Paper prepared for the DfID. University of Oxford, Oxford.

Easterly, W. (2001). "Middle Class Consensus and Economic Development", *Journal of Economic Growth*, 6(4): 317-336.

Elbers, C. & J. Gunning (2003). "Vulnerability in a Stochastic Dynamic Model", Discussion Paper 2003-070/2. Tinbergen Institute, Rotterdam.

Erikson, R., J. Goldthorpe, & L. Portocarero (1979). "Intergenerational Class Mobility in Three Western European Societies", *British Journal of Sociology*, 30(4): 415-441.

Gaiha, R. & K. Imai (2008), "Measuring Vulnerability and Poverty Estimates for Rural India", Research Paper No. 2008/40. UNU-WIDER, Helsinki.

Goldthorpe, J.H. (1987). *Social Mobility and Class Structure in Modern Britain*. Oxford: Clarendon Press.

Goldthorpe, J.H. & A. McKnight (2004). "The Economic Basis of Social Class", CASE Paper 80. London School of Economics, London.

Hertova, D., L.F. Lopez-Calva, & E. Ortiz-Juarez (2011). "Bigger… but Stronger? The Middle Class in Chile and Mexico in the Last Decade", Inclusive Development Series No. 2. RBLAC-UNDP, New York.

Hoddinott, J. & A. Quisumbing (2003). "Methods for Microeconometric Risk and Vulnerability Assessments", Social Protection Discussion Paper 0323. The World Bank, Washington D.C.

Kharas, H. & G. Gertz (2010). "The New Global Middle Class: A Cross-Over from West to East", in Cheng Li (ed.) *China's Emerging Middle Class: beyond Economic Transformation*. Washington, DC: Brookings Institution Press.

Ligon, E. & L. Schechter (2004). "Evaluating Different Approaches to Estimating Vulnerability", Social Protection Discussion Paper 0210. The World Bank, Washington D.C.

Milanovic, B. & S. Yitzhaki (2002). "Does Decomposing World Income Distribution: Does the World Have a Middle Class?" *Review of Income and Wealth*, 48(2): 155-178.

Ravallion, M. (2010). "The Developing World's Bulging (but Vulnerable) Middle Class", *World Development*, 38(4): 445-454.

Sen, A. (1983). "Poor, Relatively Speaking", *Oxford Economic Papers*, 35(2): 153-169.

Wright, E.O. (1979). *Class Structure and Income Determination*. New York: Academic Press.

Table A1: Determinants of being poor and income, Chile: 2001-2006[/a]

Logistic and linear regression analysis

Model:	Logistic		Linear	
Dependent Variable:	Poverty		Income (log-scale)	
	Coeff.	**S.E.**	**Coeff.**	**S.E.**
Education of the head	-0.249***	(0.057)	0.143***	(0.013)
Age of the head	0.064	(0.048)	0.006	(0.010)
Age squared of head	-0.001**	(0.001)	0.000	(0.000)
Sex of the head (1 = male)	-0.429	(0.275)	0.320***	(0.078)
Head without social insurance[/b]	0.367**	(0.162)	-0.195***	(0.042)
Unfinished floor	0.788***	(0.249)	-0.099	(0.072)
Household without sanitation	0.402	(0.267)	-0.319***	(0.072)
Head cohabiting (omitted)	…	…	…	…
Head married	0.030	(0.206)	-0.035	(0.051)
Head without partner	0.002	(0.291)	0.442***	(0.083)
Head in agriculture (omitted)	…	…	…	…
Head as unskilled manual worker	-0.026	(0.220)	-0.025	(0.051)
Head as skilled manual worker	-0.181	(0.253)	0.088	(0.058)
Head as independent worker	0.010	(0.234)	0.158***	(0.059)
Head in clerical activities	-0.648**	(0.329)	0.146**	(0.066)
Head as professional manager	-2.794***	(1.056)	0.427***	(0.088)
Region VII (omitted)	…	…	…	…
Region III	-0.717**	(0.341)	0.293***	(0.074)
Region VIII	0.094	(0.177)	0.018	(0.042)
Metropolitan region	-0.460**	(0.191)	0.247***	(0.041)
Rurality	-0.827***	(0.205)	-0.061	(0.046)
Ocurrence of health shocks 2001-2006	0.125	(0.164)	0.037	(0.038)
Change in number of members working 2001-2006	-0.543***	(0.094)	-0.158***	(0.017)
Change in household size 2001-2006	0.207**	(0.093)	0.164***	(0.022)
Constant	-0.306	(1.128)	9.603***	(0.248)
Observations	1,950		1,950	
Pseudo R^2 / R^2	0.138		0.385	

Source: Author's calculations based on data from CASEN Panel 2001-2006.

[/a] Dependent variables are the poverty status of households in logistic model, and the household per capita income (log-scale) in linear model. [/b] It refers to the coverage to different provisional systems.

Robust standard errors in parentheses, * p<0.05, ** p<0.01, *** p<0.001

Table A2: Determinants of being poor and income, Mexico: 2002-2005[/a]

Logistic and linear regression analysis

Model:	Logistic		Linear	
Dependent Variable:	Poverty		Income (log-scale)	
	Coeff.	**S.E.**	**Coeff.**	**S.E.**
Education of the head	-0.336***	(0.051)	0.228***	(0.017)
Age of the head	-0.018	(0.017)	-0.005	(0.006)
Age squared of head	0.000	(0.000)	0.000*	(0.000)
Sex of the head (1 = male)	-0.275*	(0.165)	0.356***	(0.068)
Head without social insurance[/b]	0.521***	(0.083)	-0.310***	(0.032)
Unfinished floor	0.720***	(0.105)	-0.348***	(0.046)
Household without sanitation	0.087	(0.138)	-0.143**	(0.067)
Head cohabiting (omitted)	…	…	…	…
Head married	-0.117	(0.112)	-0.017	(0.038)
Head without partner	-0.596***	(0.172)	0.529***	(0.067)
Head in agriculture (omitted)	…	…	…	…
Head as unskilled manual worker	-0.340**	(0.133)	0.367***	(0.050)
Head as skilled manual worker	-0.695***	(0.112)	0.383***	(0.045)
Head as independent worker	-0.541***	(0.208)	0.279***	(0.083)
Head in clerical activities	-0.987***	(0.237)	0.589***	(0.064)
Head as professional manager	-1.026***	(0.215)	0.706***	(0.065)
Head in commerce and services	-0.388***	(0.139)	0.277***	(0.058)
Head in army, police, and other	-1.007***	(0.240)	0.328***	(0.071)
South region (omitted)	…	…	…	…
Central region	-0.346***	(0.111)	0.174***	(0.043)
Western region	-0.491***	(0.115)	0.259***	(0.046)
Northwest region	-0.995***	(0.127)	0.338***	(0.042)
Northeast region	-0.411***	(0.116)	0.354***	(0.043)
Rurality	-0.017	(0.093)	-0.351***	(0.031)
Ocurrence of shocks 2001-2005	0.110	(0.086)	-0.034	(0.032)
Change in number of members working 2002-2005	-0.405***	(0.038)	-0.118***	(0.013)
Change in household size 2002-2005	0.221***	(0.038)	0.034**	(0.014)
Constant	0.871*	(0.494)	5.956***	(0.181)
Observations	4,514		4,514	
Pseudo R^2 / R^2	0.154		0.355	

Source: Author's calculations based on data from MxFLS 2002-2005.

[/a] Dependent variables are the poverty status of households in logistic model, and the household per capita income (log-scale) in linear model. [/b] It includes medical services provided by IMSS, ISSSTE, PEMEX, SEDENA, MARINA, local institutions, universities and private suppliers.

Robust standard errors in parentheses, * p<0.05, ** p<0.01, *** p<0.001

Table A3: Determinants of being poor and income, Peru: 2002-2006[/a]
Logistic and linear regression analysis

Model:	Logistic		Linear	
Dependent Variable:	Poverty		Income (log-scale)	
	Coeff.	**S.E.**	**Coeff.**	**S.E.**
Education of the head	-0.564***	(0.049)	0.245***	(0.014)
Age of the head	-0.023	(0.017)	0.011**	(0.005)
Age squared of head	0.000	(0.000)	0.000	(0.000)
Sex of the head (1 = male)	-0.060	(0.170)	0.166***	(0.053)
Head without social insurance[/b]	-0.258***	(0.094)	0.126***	(0.026)
Unfinished floor	0.620***	(0.093)	-0.262***	(0.027)
Household without sanitation	0.382***	(0.103)	-0.147***	(0.031)
Head cohabiting (omitted)	…	…	…	…
Head married	-0.238**	(0.102)	0.093***	(0.027)
Head without partner	-0.857***	(0.159)	0.402***	(0.050)
Head in agriculture (omitted)	…	…	…	…
Head in mining, electricity, gas and water	-0.648*	(0.333)	0.988***	(0.110)
Head in manufacturing	-0.695***	(0.190)	0.366***	(0.049)
Head in construction	-0.302	(0.220)	0.294***	(0.054)
Head in commerce	-0.496***	(0.138)	0.424***	(0.039)
Head in transport and communications	-0.498***	(0.192)	0.361***	(0.047)
Head in government and clerical activities	-0.846***	(0.197)	0.547***	(0.048)
Head in other services	-0.729***	(0.197)	0.342***	(0.049)
Selva region (omitted)	…	…	…	…
North Coast	-0.322**	(0.127)	0.191***	(0.035)
Central Coast	-1.072***	(0.206)	0.219***	(0.046)
Southern Coast	-0.518**	(0.232)	0.295***	(0.057)
Northern Sierra	0.466**	(0.193)	-0.080	(0.062)
Central Sierra	0.578***	(0.134)	-0.121***	(0.039)
Southern Sierra	0.292**	(0.145)	-0.065	(0.041)
Metropolitan area of Lima	-0.198	(0.167)	0.288***	(0.043)
Rurality	-0.048	(0.122)	-0.073**	(0.030)
Ocurrence of health shocks 2002-2006	-0.093	(0.088)	0.003	(0.025)
Change in number of members working 2002-2006	-0.090**	(0.039)	-0.105***	(0.010)
Change in household size 2002-2006	0.218***	(0.028)	0.077***	(0.007)
Constant	2.533***	(0.502)	3.762***	(0.145)
Observations	3,258		3,258	
Pseudo R^2 / R^2	0.194		0.486	

Source: Author's calculations based on data from ENAHO Panel 2002-2006.
[/a] Dependent variables are the poverty status of households in logistic model, and the household per capita income (log-scale) in linear model. [/b] It includes medical services provided by ESSALUD, private suppliers, *Entidad prestadora de salud*, insurance from FF.AA., *Seguro Integral de Salud*, universities, and other suppliers.
Robust standard errors in parentheses, * p<0.05, ** p<0.01, *** p<0.001

Table A4: Specifications in robustness analysis

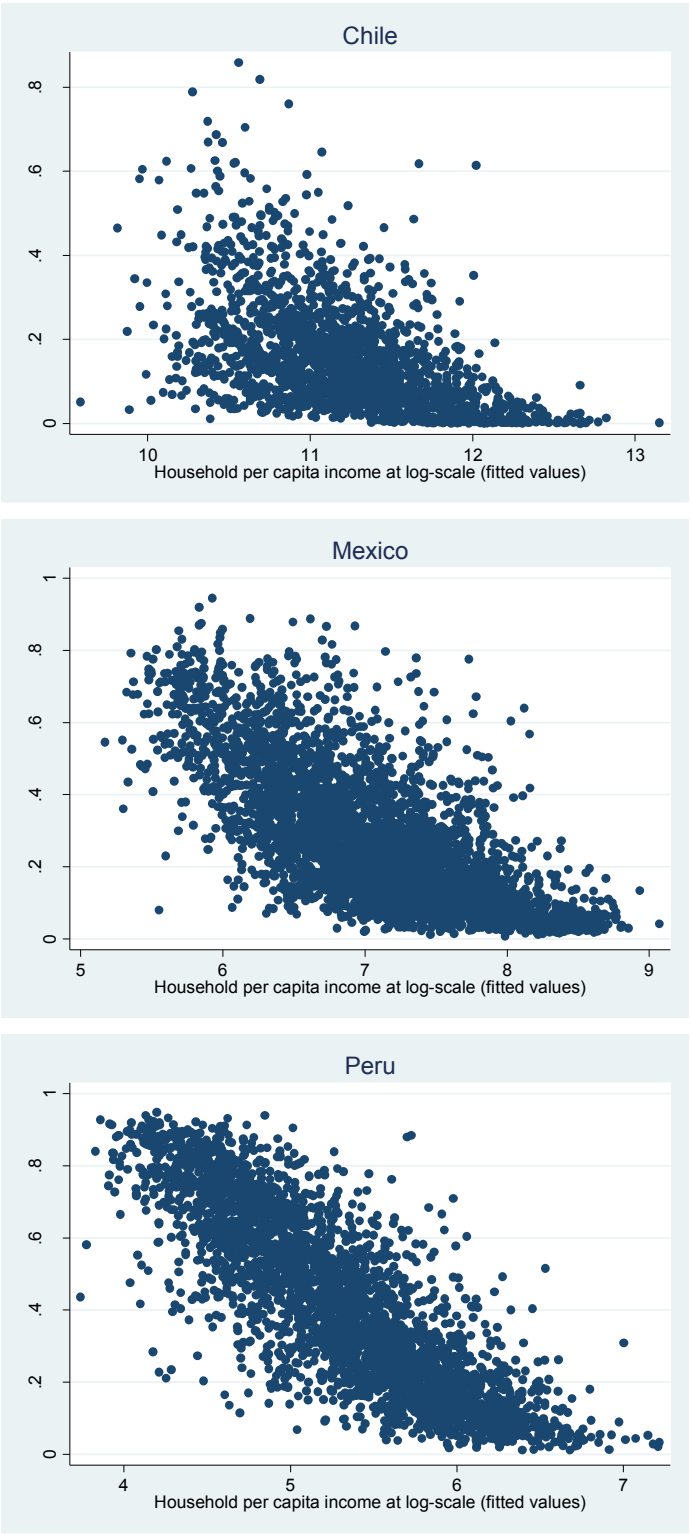
Chile	Mexico	Peru
Education of the head	Education of the head	Education of the head
Age of the head	Age of the head	Age of the head
Age squared of head	Age squared of head	Age squared of head
Sex of the head (1 = male)	Sex of the head (1 = male)	Sex of the head (1 = male)
Head without social insurance	Head without social insurance	Head without social insurance
Unfinished floor	Unfinished floor	Unfinished floor
Household without sanitation	Household without sanitation	Household without sanitation
Head cohabiting (omitted)	Head cohabiting (omitted)	Head cohabiting (omitted)
Head married	Head married	Head married
Head without partner	Head without partner	Head without partner
Head in agriculture (omitted)	Head in agriculture (omitted)	Head in agriculture (omitted)
Head as unskilled manual worker	Head as unskilled manual worker	Head in minning, electricity, gas and water
Head as skilled manual worker	Head as skilled manual worker	Head in manufacturing
Head as independent worker	Head as independent worker	Head in construction
Head in clerical activities	Head in clerical activities	Head in commerce
Head as professional manager	Head as professional manager	Head in transport and communications
Region VII (omitted)	Head in commerce and services	Head in government and clerical activities
Region III	Head in army, police, and other	Head in other services
Region VIII	South region (omitted)	Selva region (omitted)
Metropolitan region	Central region	North Coast
Rurality	Western region	Central Coast
Ocurrence of health shocks 2001-2006	Northwest region	Southern Coast
Change in # of members working 2001-2006	Northeast region	Northern Sierra
Change in household size 2001-2006	Rurality	Central Sierra
	Ocurrence of shocks 2001-2005	Southern Sierra
	Change in # of members working 2002-2005	Metropolitan area of Lima
	Change in household size 2002-2005	Rurality
		Ocurrence of health shocks 2002-2006
		Change in # of members working 2002-2006
		Change in household size 2002-2006

Table A5: Two-group mean-comparison *t* test

	Significance level (95%) Pr(|T| > |t|)		Average by groups		
			Lower Class		Middle class
Chile	PL-$10/$10-50	<PL/$10-50	<PL	PL-$10	$10-50
No education	0.492	0.170	4.9	7.3	6.6
Incomplete primary	0.025	0.000	47.5	38.1	33.7
Complete primary	0.050	0.000	18.9	12.9	10.3
Complete secondary	0.996	0.000	8.8	16.8	16.8
Tertiary education	0.000	0.000	0.9	1.9	9.0
No insurance	0.271	0.000	21.2	15.2	13.6
Dirt floor	0.000	0.000	10.7	6.4	3.2
No sanitation	0.000	0.000	8.5	4.7	1.3
Head is farmer	0.000	0.000	40.3	23.4	12.9
Head in clerical activities	0.008	0.000	5.4	9.6	13.8
Head is professional/manager	0.000	0.000	0.9	3.9	13.1
Rurality	0.000	0.000	33.8	22.7	13.5
Ocurrence of shocks	0.000	0.000	24.1	26.5	33.7
Mexico	PL-$10/$10-50	<PL/$10-50	<PL	PL-$10	$10-50
No education	0.000	0.000	20.5	11.1	6.9
Incomplete primary	0.000	0.000	36.8	27.5	18.9
Complete primary	0.029	0.007	22.5	21.8	19.0
Complete secondary	0.000	0.000	1.9	4.6	7.6
Tertiary education	0.000	0.000	2.2	6.2	20.5
No insurance	0.000	0.000	64.8	39.3	23.8
Dirt floor	0.000	0.000	24.8	10.0	3.5
No sanitation	0.000	0.000	11.6	5.0	2.0
Head is farmer	0.000	0.000	48.4	19.5	10.3
Head in clerical activities	0.000	0.000	1.2	5.9	9.8
Head is professional/manager	0.000	0.000	2.3	5.8	19.0
Rurality	0.000	0.000	76.8	48.5	31.5
Ocurrence of shocks	0.003	0.003	28.4	28.4	24.2
Peru	PL-$10/$10-50	<PL/$10-50	<PL	PL-$10	$10-50
No education	0.000	0.000	13.0	7.6	2.7
Incomplete primary	0.000	0.000	34.9	21.8	11.3
Complete primary	0.152	0.000	21.3	15.7	13.3
Complete secondary	0.190	0.000	12.1	23.1	25.7
Tertiary education	0.000	0.000	0.5	5.1	15.2
No insurance	0.100	0.948	28.8	32.2	28.7
Dirt floor	0.000	0.000	69.4	34.2	17.4
No sanitation	0.000	0.000	35.9	15.0	6.9
Head is farmer	0.000	0.000	71.6	31.5	16.6
Head in clerical activities	0.001	0.000	2.4	9.6	15.0
Head is professional/manager	…	…	…	…	…
Rurality	0.000	0.000	77.5	44.8	26.8
Ocurrence of shocks	0.268	0.106	48.0	49.1	51.7

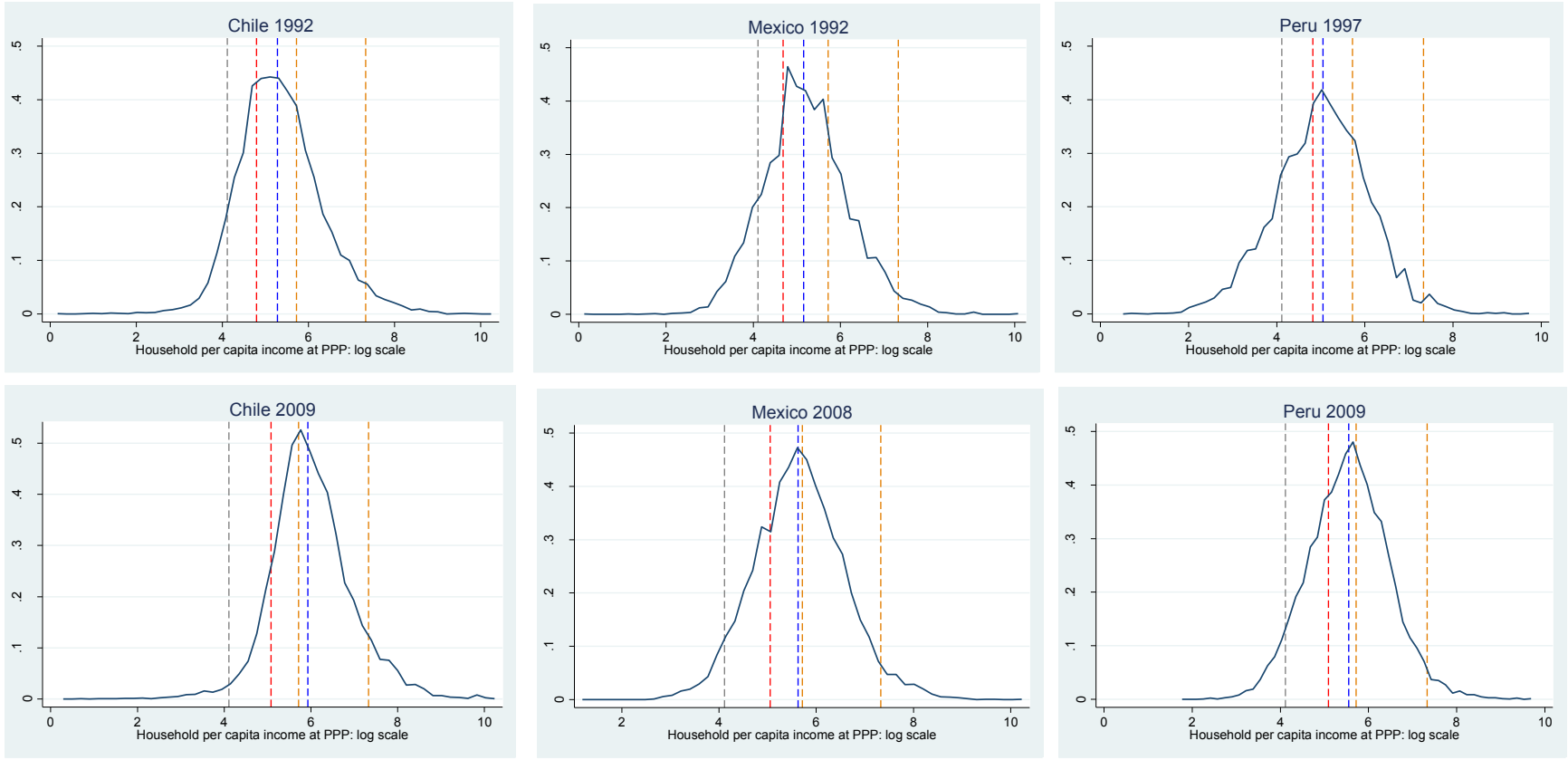
Source: Author's calculations based on data from CASEN, MxFLS and ENAHO panel databases.
The difference in the occurrence of shocks was found to be statistically significant in Chile and Mexico; however, the incidence is adverse for the middle class in the former. This could be due to the different definitions of shocks: while in Chile (and in Peru) it is only possible to identify health shocks, in Mexico the variable includes a broader range of events.

Figure A1: Correlation plots between the estimated probabilities of fall into poverty and income (fitted values)



Source: Author's calculations based on data from CASEN, MxFLS and ENAHO panel databases.

Figure A2: Kernel distributions of income in Chile, Mexico and Peru



Source: Author's calculations based on data from CASEN, ENIGH and ENAHO cross-sectional databases.
/a The red line represent the official poverty line; the blue line is the median income; the gray lines is the lower threshold of Banerjee and B&D; and the yellow lines are the thresholds proposed in this paper.