LSE Department of Social Policy

Working Paper
January 2024

Degrees of vulnerability to poverty:

A low-income dynamics approach for

Chile

Joaquín Prieto-Suarez

Department of Social Policy
London School of Economics and
Political Science Houghton Street
London WC2A 2AE

Email: soc.pol.webteam@lse.ac.uk

Telephone: +44 (0)20 7955 6001

lse.ac.uk/social-policy

To cite this paper: Prieto-Suarez, J. (2024) Degrees of vulnerability to poverty: A low-income dynamics approach for Chile, Social Policy Working Paper 01-24, London: LSE Department of Social Policy.

Degrees of vulnerability to poverty: A low-income dynamics approach for Chile[1]

Joaquín Prieto[2]

Abstract

I propose an empirical framework to identify different degrees of vulnerability to poverty using two vulnerability lines that classify currently non-poor people into risk groups: high, moderate and low risk of falling into poverty in the next period. The latter corresponds to the income secure middle class. My approach makes two contributions. First, it extends recent research that defines the middle class using a vulnerability threshold by introducing a new subdivision of the vulnerable group that would be useful in practice for public policy objectives. Second, it uses two models to predict both the probability of entering poverty and household income as part of the estimation procedures. The former controls for initial conditions effects and attrition bias, and the latter addresses the retransformation problem. I apply my approach to Chile using longitudinal data from the P-CASEN 2006-2009. The resulting vulnerability cut-offs (using the upper-middle-income country poverty line) are \$20.0 per person per day for the low vulnerability line and \$9.9 pppd for the high vulnerability line (both in 2011 PPP). My vulnerability lines differ significantly from those estimated in previous research on vulnerability and the middle class in Latin America. I argue that previous research has underestimated the size of the population at risk of falling into poverty and overestimated the growth of the middle class. Misclassifying the vulnerable as middle class limits their access to anti-poverty policies.

Keywords: Chile; Latin America; longitudinal data; middle class; poverty dynamics; vulnerability to poverty

JEL Classification: C18; D31; D63; I32

---

[1] I would like to thank Stephen Jenkins, Grace Lordan, Frank Cowell and Philippe Van Kerm for their comprehensive comments. I would also like to acknowledge the feedback from the participants of the Eighth ECINEQ Meeting and the ISA-RC28 Conference 2019. I am grateful for the support of CONICYT's Becas Chile programme (CONICYT/PFCHA/DOCTORADO BECAS CHILE/2015-72150404).

[2] International Inequalities Institute at the LSE, prietosj@lse.ac.uk

1. Introduction

Over the past decade, international agencies and some governments in developing countries have adopted a new forward-looking perspective in social policy design, identifying those who are not poor but at risk of falling into poverty. See Birdsall et al. (2014) for Latin America, Klasen & Waibel (2015) for South-East Asia, and Dang & Dabalen (2018) for Africa. Knowing ex ante which households are vulnerable to poverty makes it possible to design effective anti-poverty strategies and improve risk management policies, such as risk insurance programmes and incentives for self-protective savings (Dercon, 2005). However, although the concept of vulnerability to poverty emerged in the mid-1990s (Morduch, 1994; Ravallion, 1996), there is still no consensus on how to define and measure it due to the difficulty of analysing unknown future distributions of poverty (Ceriani, 2018; Gallardo, 2018).

One of the challenges in identifying those vulnerable to poverty in a society is that within this group, households face different levels of vulnerability (Lipton, 1983). Although there is an extensive economic literature that provides methods to divide vulnerable people into groups according to their level of vulnerability,[3] this issue has not been addressed in the literature that defines the middle class using a vulnerability threshold (Dang & Lanjouw, 2017; López-Calva & Ortiz-Juarez, 2014; Schotte, Zizzamia, & Leibbrandt, 2018). This approach, which focuses on the definition of the middle class, has been used by international agencies and governments to design social protection policies and monitor groups at risk of poverty in regions such as Latin America (e.g., Ferreira et al., 2013; Stampini, Robles, Sáenz, Ibarrarán, & Medellín, 2016).[4] While it is true that using poverty and vulnerability lines is a simple way to classify the population into poor, vulnerable and middle class, the problem is that a household close to the poverty line is very different from a household just below a vulnerability line that identifies the economically secure middle class.

Following the literature that classifies social groups with different degrees of vulnerability to poverty, my main aim in this study is to identify three groups: the non-poor with a low, medium and high probability of falling into poverty. I define vulnerability to poverty as the risk of the non-poor in the current year falling into poverty in the following year, based on the approach that

[3] Examples of vulnerable groups are: high vulnerability and low vulnerability (Suryahadi & Sumarto, 2003), poverty-induced vulnerability and risk-induced vulnerability (Günther & Harttgen, 2009), moderate vulnerability and severe vulnerability (Gallardo, 2013) and relative vulnerability and high vulnerability (Feeny & McDonald, 2016).

[4] In April 2018, the World Bank updated the vulnerability line for upper-middle-income countries from $10.0 dollars pppd in 2005 PPP (this cut-off updates López-Calva & Ortiz-Juarez (2014) work) to $13.0 dollars pppd 2011 PPP. More information can be found in the following link http://www.worldbank.org/en/topic/poverty/lac-equity-lab1/poverty/head-count

considers vulnerability as expected poverty (e.g. Chaudhuri (2003) and Christiaensen & Subbarao (2005)).

My research presents a methodological contribution in the literature on poverty vulnerability. Essentially, I extend the approach of Dang and Lanjouw (2017) in three aspects. First, I introduce a new vulnerability threshold. While their approach allows the calculation of only one vulnerability line to identify a vulnerable subpopulation, my approach can be extended to derive more than one vulnerability line and thus also identify subpopulations with different levels of vulnerability. Second, I implement an improved empirical strategy, combining elements from both the work of López-Calva and Ortiz-Juarez (2014) and the work on low-income transition models (Cappellary and Jenkins, 2004; Schote et al., 2018).[5] Lastly, in contrast to Dang and Lanjouw's (2017) strategy, the new method I propose in this research first calculate the vulnerability line, and subsequently determines the proportions of the transition matrix.

My approach follows a three-step strategy. First, I estimate the probability that a currently non-poor household will become poor in the next period. Unlike López-Calva & Ortiz-Juarez (2014), who assume a logistic model to quantify the predicted household risk of poverty, I use the first-order endogenous switching Markov model developed by Cappellari & Jenkins (2004) to estimate poverty entries for non-poor individuals. This model, which is also used by Schotte et al. (2018), allows one to simultaneously control for the potential endogeneity of unobserved heterogeneity, attrition, and initial conditions. Second, I use a log-linear model between household income and household characteristics to predict household income. To address the problem of the retransformation scale of household income (Santos Silva & Tenreyro, 2006), I follow the method proposed by Duan (1983). Third, I define the vulnerability line as the average predicted income of households whose probability of falling into poverty is within $\pm$ 1 percentage point of the poverty entry rate estimated for the non-poor population.

Specifically, I estimate two lines of vulnerability. The first is a high vulnerability line, which focuses on households in the central part of the income distribution. The second is a low vulnerability line, which focuses on the upper part of the income distribution. The former identifies two groups: those at high risk and those at moderate risk of falling into poverty in the next period. The latter identifies the third group: those with a low risk of falling into poverty. The second vulnerability

[5] My approach builds on three papers that have made significant contributions to the study of vulnerability to poverty: Dang & Lanjouw (2017), who derive the income cut-off (vulnerability line) for India, the US and Vietnam using a non-parametric approach; López-Calva & Ortiz-Juarez (2014), who estimate a vulnerability line based on household characteristics in three countries in Latin America; and Schotte et al. (2018), who use a poverty dynamic approach to identify the (non-poor) vulnerable group in South Africa.

line therefore serves the dual purpose of identifying the lower income threshold for the secure middle class and the higher income threshold for the moderately vulnerable. Identifying these three groups on the basis of their level of vulnerability makes it possible to design policy strategies tailored to each of them. This is particularly important in countries that have managed to reduce absolute poverty but have high income mobility, which can be explained by a precarious and unstable labour market and weak social safety net systems that fail to help households cope with idiosyncratic shocks (OECD, 2018a; Torche, 2005).

For policy makers, classifying non-poor households according to their degree of vulnerability has three advantages. First, an approach using two - or more - vulnerability lines can reduce the likelihood of misclassification. If a vulnerability line is associated with a single average risk of falling into poverty, this could lead to households at moderate risk being inconsistently classified as vulnerable or middle class. Instead, the use of multiple lines allows for a more nuanced classification, reducing the scope for error. Second, it offers the potential for more effective allocation of public resources. It refines the targeting of a cash transfer programme to ensure that only the most vulnerable households benefit - rather than households that are vulnerable but at lower risk of being poor - thereby increasing the welfare impact and reducing inefficiencies in the distribution of resources. Third, it identifies households that are not at the highest risk of falling into poverty but are still economically insecure. Making this group visible allows policymakers to discuss what it means for these households to be economically secure and how to ensure that policies enable them to be so.

Using my approach, I derive two vulnerability lines for Chile using four waves of panel data from the CASEN survey, covering the period between 2006 and 2009.[6] The Panel CASEN is a national household survey that, despite being a short panel (four waves), provides annual information on household income as well as on education, health, labour market, housing, and social benefits.

The income threshold I estimate to identify an income-secure middle class differs significantly from the World Bank's proposed threshold for measuring the middle class in upper-middle income countries, which is $13 pppd (2011 PPP). I estimate a threshold of $20.0 pppd (2011 PPP) for the low vulnerability line (middle class) and $9.9 pppd (2011 PPP) for the high vulnerability line. The use of my two vulnerability lines allows the discussion on the design and targeting of anti-poverty policies in Chile to focus on the current distinction between vulnerable and non-vulnerable.

[6] This Panel CASEN replaced the 'old' Panel CASEN 2001-2006 used by López-Calva & Ortiz-Juarez (2014), which collected longitudinal data over a five-year interval from a sample representative of 4 of the country's 15 regions. The 'new' Panel CASEN 2006-2009 was designed and implemented by the Chilean Ministry of Planning and the Social Observatory of the Alberto Hurtado University. For more details see OSUAH (2011a).

This paper is organised as follows. In section 2, I review the literature on vulnerability to poverty and middle-class identification and discuss the main approaches to calculating vulnerability lines. In section 3, I explain how I identify degrees of vulnerability to poverty. In Section 4, I describe data and definitions. Also, I present the descriptive statistics of poverty dynamics. In section 5, I apply my approach to the case of Chile. In Section 6, I present the conclusions.

2. Vulnerability-to-poverty and middle-income class identification: from divergent to convergent approaches

Until recently, economic research focused on the middle class (e.g., Atkinson & Brandolini (2013)) has progressed in parallel with research on the vulnerable (e.g., Chaudhuri et al., (2002)), with no relevant connection or dialogue, even though both studies are two sides of the same coin.

Measures to identify the middle class: income as the main indicator

Research focusing on the definition of the middle-class uses economic resources as the main indicator, in particular household income (Gornick & Jäntti, 2014). Indeed, the middle class is usually analysed as the middle group within the income distribution, for which several strategies have been implemented to define income thresholds, either relative or absolute.

Relative measures define the middle class by using household income to find a threshold that is anchored to the information provided by each country's income distribution. See Estache & Leipziger (2009) and Atkinson & Brandolini (2013).[7] However, these measures do not adequately compare the middle class across countries with different income distributions. In developing countries, the incomes of the middle class are much more modest than those of the middle class in developed countries. Only a minority of the population in low- and middle-income economies can be considered middle class if the economic welfare of developed countries is used as a reference (Milanovic & Yitzhaki, 2002; Ravallion, 2010).

In contrast to relative measures, absolute measures of the middle class use thresholds based on a given level of income or expenditure. Early research suggested that the lower boundary of the middle class was $2 pppd and the upper boundary was $10 or $13 pppd (Banerjee & Duflo, 2008;

[7] Among these measures, there are three main definitions: i) a distance from the median income, e.g. those whose income falls between 75 and 125 per cent of the median income are considered middle class (Birdsall, Graham, & Pettinato, 2000; Davis & Huston, 1992); ii) a range in the income distribution, e.g. those whose income falls within the 3rd and 4th quintile groups are considered middle class (Alesina & Perotti, 1996; Barro, 2000; Easterly, 2001); and iii) a certain distance from the poverty line, e.g. those whose income is above 130 per cent of the country's official poverty line are considered middle class (World Bank, 2012).

Ravallion, 2010).[8] The income cut-off used by these authors to define the middle class was highly controversial, as the vulnerable group above the $2 dollars income cut-off lacked the core characteristics of the middle class, namely income stability, access to social security benefits and contribution to the social security system through tax payments (Birdsall, 2015). In recent years, the absolute purchasing power approach has been highlighted as a strategy to compare the middle class between different countries at the global level (e.g. $11 to $110 pppd in 2011 purchasing power parity (PPP) terms) (Kharas, 2017).

Vulnerability-to-poverty approach: measuring downward mobility

In the early 2000s, the economic literature focusing on vulnerable groups developed a conceptual framework known as vulnerability to poverty (Hoddinott & Quisumbing, 2010). This literature can be divided into four groups: i) papers that emphasise the element of expected poverty, i.e. that consider vulnerability to be the probability of a household falling into poverty in a future period (e.g. Pritchett et al., (2000); Chaudhuri et al., (2002)); ii) papers that stress the element of exposure to risk, for example to indicate retrospectively whether an observed economic shock led to a loss of well-being in a household (e.g. Skoufias & Quisumbing (2005)); iii) papers that define vulnerability as the difference between a household's utility derived from certainty-equivalent consumption and its expected utility derived from actual consumption (e.g. Ligon & Schechter (2003)); and iv) papers that identify vulnerable individuals according to the expected value and a risk parameter, known as the mean risk-based approach (Chiwaula, Witt, & Waibel, 2011; Gallardo, 2013). Over the last decade, the literature on vulnerability to poverty has continued to develop around these four groups (e.g., Celidoni, 2015; Günther & Maier, 2014; Hohberg, Landau, Kneib, Klasen, & Zucchini, 2018; Klasen & Waibel, 2015), with the addition of attempts to measure vulnerability from a multidimensional perspective (e.g., Feeny & McDonald, 2016; Gallardo, 2019).

The most used of these approaches is vulnerability as expected poverty (VEP). It has the advantage of being relatively simple to implement with widely available or easily collected data, as well as being a forward-looking concept that is easier for policymakers to understand and interpret than the other three definitions (Hohberg et al., 2018).

Using the vulnerability line to identify both the vulnerable and the middle class

While the economic literature has used income thresholds to identify the middle class, most VEP studies define vulnerability thresholds in terms of a certain probability of falling into poverty.

[8] The lower threshold corresponds to the World Bank's poverty line for developing countries, and the upper threshold of $13 proposed by Ravallion (2010) corresponds to the poverty line in the United States.

However, three new methods based on the VEP approach have linked the vulnerability threshold (risk) to household income levels, namely those developed by López-Calva & Ortiz-Juarez (2014); Dang & Lanjouw (2017); and Schotte et al. (2018).[9] This income threshold, known as the vulnerability line, allows the identification of the vulnerable and non-vulnerable in the same way as economic research on the middle class, as discussed above. Households are considered vulnerable to poverty if their income is below the vulnerability line, and middle class if their income is just above the vulnerability line.

The vulnerability line is defined as the income V_t such that having an income y_t below V_t at t (but above the poverty line Z at t) means that the risk of being poor at t+1 ($\Pr\,(y_{t+1} < Z)$) is greater than or equal to some critical probability level, known as the risk threshold. The vulnerability line V_t distinguishes households that are still vulnerable to poverty from those groups that are economically more secure. This definition is closely related to the Weberian notion that households should enjoy a certain minimum of economic security to be considered middle-class (Goldthorpe & McKnight, 2006; López-Calva & Ortiz-Juarez, 2014). This vulnerability line has served the purpose of closing the gap between the research on the income-secure middle class and the vulnerable group (Schotte et al., 2018).

López-Calva and Ortiz-Juarez's study (2014) introduces a vulnerability line connected with household characteristics extracted from longitudinal household surveys from Chile, Mexico, and Peru. Their model consists of two parts: a logistic model estimating the poverty probability based on household characteristics for a non-poor sample, and a log-linear regression model predicting per capita income from the same variables. The researchers set a 10% poverty risk as the cut-off between economic stability and vulnerability, establishing the forecasted income at this probability as the reference point demarcating the bottom threshold of the middle class. One of the main advantages of the López-Calva & Ortiz-Juarez (2014) approach is that it provides a vulnerability line that can be used to compare upper-middle-income countries since it is based in the World Bank poverty line for these countries. This explains its extensive use to identify and measure those who are vulnerable to poverty and also those who qualify as middle-class in contexts of absolute poverty reduction (e.g. Birdsall et al., 2014; Ferreira et al., 2013; Stampini et al., 2016; Wietzke & Sumner, 2018).

Schotte et al. (2018) use the observed average rate of poverty entry for the non-poor population as a probability cut-off to separate the vulnerable from the middle class in South Africa. They calculate

[9] These methods assume that there is a monotonic relationship between the predicted poverty entry probability and the household income. Although this assumption is plausible, it does not guarantee that higher base period income (among the non-poor) implies a lower probability of falling into poverty.

the vulnerability line as "the average monthly per capita household expenditure of those respondents whose predicted poverty transition probability falls within the 95 percent confidence interval around [this] probability threshold". Importantly, they use the Cappellari and Jenkins (2004) poverty dynamics model to estimate the poverty entry probability for non-poor people. This model estimates poverty transitions probabilities while simultaneously controlling for attrition and for initial condition effects.[10] Schotte et al. (2018) use the probability cut-off and a vulnerability line to distinguish between those who are not poor, but vulnerable, and those who are middle class. Their results show significant misclassification, with 40% of households identified as vulnerable based on their observed income misclassified by poverty risk - in fact, they would more appropriately be classified as middle class. Similarly, 20% of those identified as middle class based on their observed income position would be considered vulnerable when analysed through the lens of poverty risk.[11]

Finally, the approach of Dang & Lanjouw (2017) differs from that of López-Calva & Ortiz-Juarez (2014) and Schotte et al. (2018) in that they use a non-parametric estimation method to estimate vulnerability lines as a function of household consumption or income. Thus, their approach does not use information on household characteristics. Dang & Lanjouw (2017) derive vulnerability lines that allow them to distinguish the population that is not currently poor, but vulnerable to poverty. They define the vulnerability line as the income threshold at which a given share of the population with a consumption level above this line in a given year will fall below the poverty line in the following year. They also propose a second definition that focuses on those with a consumption level above the poverty line but still below the at-risk threshold in the given year. The main advantages of Dang & Lanjouw's (2017) approach are: i) unlike studies that fix the vulnerability index at 50 percent (e.g., Chiwaula et al. (2011)), their vulnerability index is flexible; it can change or adapt based on practical complexities related to the design of social programmes, such as budget planning or targeting issues; and ii) the implementation of this approach is simple and intuitively understandable for policymakers.

3. A low-income dynamics approach to identify degrees of vulnerability to poverty

In this section, I discuss the three steps I take to determine the degree of vulnerability to poverty. First, I explain the econometric approach to modelling poverty transition probabilities. Second, I describe my proposal to derive a vulnerability line from the poverty entry rates of the non-poor in

[10] Whether a household is poor or non-poor in the base year could be a non-random event.

[11] Yet, the authors do not mention that their results show that the poverty dynamic model used for estimating the entry probabilities is not a guarantee of a monotonic relationship between income and the predicted entry probability.

the base year, and third, I show how to extend my approach to have two vulnerability lines rather than one.

First step: A first-order Markov approach to modelling poverty entries

In the initial step, I employ the endogenous switching model proposed by Cappellari & Jenkins (2004) to identify the relationship between household characteristics at t and poverty transitions probabilities, and specifically the probability of falling into poverty between t and $t+1$ for non-poor people. This model is a Markovian transition model approach and provides estimates that address two important sources of bias.[12]

First, there is the bias that arises from ignoring the problem of initial conditions. This refers to the fact that the group who are poor in the base period may be a non-random sample of the population. Ignoring this may bias poverty transition estimates because it is difficult to assume that being poor in the base year is exogenous and uncorrelated with unobserved characteristics (Jenkins, 2011). For example, unobservables can make individuals more likely to be at the lowest extreme of income distribution in a given year. Second, there is potential bias resulting from non-random survey attrition. If the attrition process is not random and is correlated with the probability of poverty entry, estimates of the relation between poverty entries and covariables may be biased as a result of endogenous selection. For example, individuals that are more likely to be observed successively in the panel can be less likely to fall into poverty compared to those that attrit.[13]

In order to address the initial conditions problem and non-random panel attrition, I employ the Cappellari & Jenkins (2004) model. The model accounts for the endogeneity of both processes to poverty transitions probabilities by freely estimating the correlations between unobservables affecting. Thus, the model consists of three equations: i) an equation for the poverty status in the base year t (in order to account for the initial conditions problem); ii) an equation for sample retention from one wave to the next (to account for non-random attrition bias); and iii) the main equation of interest for conditional poverty status in year $t+1$ for all of the pooled annual transitions.

The latent propensities for these equations are represented by $P_{i,t}^{*}$ (poverty status in the base period t), $R_{i,t+1}^{*}$ (retention in the sample between t and $t+1$), and $P_{i,t+1}^{*}$ (conditional poverty status in

[12] See Jenkins (2011) for a detailed review of the standard approaches used to model poverty transitions such as hazard regression models, covariance structure models and variance component models.

[13] López-Calva & Ortiz-Juarez (2014) model assumes a logit relationship between the poverty entry probability for the non-poor and observable variables without taking into consideration panel attrition, which is significant for the data they used. In the case of Chile, the panel data at $t+s$ analysed by López-Calva & Ortiz-Juarez (2014) is non-randomly selected (Bendezú, Denis, & Zubizarreta, 2007), and it biases estimates of some measures such as income mobility (Paredes, Prieto, & Zubizarreta, 2006).

period $t+1$), and modelled using the following linear specifications:

$$P_{i,t}^{*}=\beta' Z_{i,t}+v_{i,t} \qquad \text{with} \quad v_{i,t}=o_i+\pi_{i,t}\sim N(0,1) \tag{1}$$

$$R_{i,t+1}^{*}=\psi' W_{i,t}+\varepsilon_{i,t+1} \qquad \text{with} \quad \varepsilon_{i,t+1}=\eta_i+\xi_{i,t+1}\sim N(0,1) \tag{2}$$

$$P_{i,t+1}^{*}=\left[(P_{i,t})\gamma_1'+\left(1-P_{i,t}\right)\gamma_2'\right]X_{i,t}+u_{i,t+1} \qquad \text{with} \quad u_{i,t+1}=\mu_i+\delta_{i,t+1}\sim N(0,1) \tag{3}$$

where $i=1,\dots,N$ indexes individuals and $Z_{i,t}$, $W_{i,t}$ and $\mathrm{X}_{i,t}$ are vectors of explanatory variables characterising individual i in her household in terms of base year vales, β, ψ, γ_1 and γ_2 are vectors of parameter, and $v_{i,t}$, $\varepsilon_{i,t+1}$, and $u_{i,t+1}$ are the error terms defined as the sum of a normal individual-specific effect ($o_i, \eta_i, u_{i,t+1}$) plus a normal orthogonal white noise error ($\pi_{i,t}$, $\xi_{i,t+1}$, $\delta_{i,t+1}$) where the latter follows a standard normal distribution.

In equation (1), an individual is poor if her latent poverty propensity is greater than 0. In equation (2), if an individual's latent propensity is > 0, I observe her income in period $t+1$, otherwise I do not. If I do not observe the household's income in period $t+1$, then it is not possible to determine an individual's poverty status in period $t+1$, nor any poverty transition that may have taken place. In equation (3) I assume that an individual is in poverty in period $t+1$ if her latent propensity exceeds 0. Similar to equation 2, I can only observe $P_{i,t+1}^{*}$ if an individual does not drop out of the survey and we observe her household income in period $t+1$.

I estimate the model assuming that the joint distribution of these error terms is trivariate standard normal. The unobserved heterogeneity, that is, the individual-specific component of the error term, can be summarised by the following three correlation coefficients:

$$\rho_1 \equiv \mathrm{corr}\left(u_{i,t+1}, v_{i,t}\right)=\mathrm{cov}(\mu_i, o_i) \tag{4}$$

$$\rho_2 \equiv \mathrm{corr}\left(v_{i,t}, \varepsilon_{i,t+1}\right)=\mathrm{cov}(o_i, \eta_i) \tag{5}$$

$$\rho_3 \equiv \mathrm{corr}\left(u_{i,t+1}, \varepsilon_{i,t+1}\right)=\mathrm{cov}(\mu_i, \eta_i) \tag{6}$$

The identification of the correlation coefficients requires exclusion restrictions. Therefore, in order to allow the identification of equations (1), (2) and (3), Cappellari and Jenkins (2004) suggest using instrumental variables for both endogenous selection mechanisms that are correlated with the initial poverty status and with the attrition of the sample in the base year (t) but that are not correlated with the poverty status in time $t+1$.[14]

[14] $\mathrm{X}_{i,t}$ is a vector of covariates that has an impact on the conditional poverty status in the next period $t+1$. The vector of covariates for the initial poverty equation $\mathrm{Z}_{i,t}$ is the same as $\mathrm{X}_{i,t}$ with additional exclusion restrictions, and similarly, $\mathrm{W}_{i,t}$ is vector of the variables that determine retention, including those in $\mathrm{X}_{i,t}$, plus a number of exclusion restrictions. The inclusion of a retention equation allows for using an unbalanced panel and therefore for drawing on all the

Following other studies (e.g. Cappellari and Jenkins (2004), Ayllon (2013), and Schotte et al. (2018)) I use two types of instrumental variables. First, as an instrumental variable for the retention of the sample I use a dichotomous variable that identifies, among all the survey respondents, those individuals who were original members of the sample (interviewed in the first round), distinguishing them from those temporarily integrated into the panel sample because they were part of a household with an original member. The rationale behind this variable is that the original sample members have a higher probability of continuing in the sample than the temporary members regardless of the income level of their households.

Second, I use retrospective recall data as instrumental variables for the initial condition of poverty: i) the levels of education of the mother and father of each respondent; as well as ii) the type of work of both parents. The assumption behind these variables is that both the level of education of the parents and the work they did in the past affects the initial condition of poverty in the base year for the individual that belongs to the panel sample but does not directly affect transitions of poverty of the individual from one year to another.

Using the estimated parameter values of my model, I derive the probability of entering poverty for each non-poor household in the base year (t). This probability is the proportion of households that are non-poor in period t that become poor in $t+1$. Specifically, the poverty entry probability ($e_{i,t+1}$) can be written as follows:

$$e_{i,t+1} = \Pr\left(P_{i,t+1} = 1 | P_{i,t} = 0\right) = \frac{\Phi_2(\gamma_2' x_{i,t}; -\beta' z_{i,t}; -\rho_2)}{\Phi(-\beta' z_{i,t})} \tag{7}$$

where $\Phi_2(\cdot)$ and $\Phi(\cdot)$ denote respectively the cumulative density functions of the trivariate and bivariate standard normal distribution (for details refer to section 2 in Cappellari & Jenkins, 2004).

Finally, I estimate the poverty entry rate between t and $t+1$ ($\bar{e}_{i,t+1}$) as the average probability ($\mathrm{N}^{-1} \sum_{i=1}^{N} \widehat{e_{i,t+1}}$) of falling into poverty for a non-poor household.

Cappellari and Jenkins's model (2004) has two additional advantages related to the use of panel data. It can be applied to relatively short panels because it only requires two waves of data, and the model can accommodate left-censored poverty spells because of its first-order Markov assumption. Individuals who remain in the same state at each wave (i.e. are always poor or never poor) are included in the estimation sample.[15]

information available in the panel.

[15] Other poverty transition models without a first-order Markov assumption such as hazard models can control for duration dependence. But the price paid is they cannot accommodate left-censored poverty spells. This may bias

Second step: Strategy to associate predicted poverty entry rates with a household's per capita income level

The predicted poverty entry rate ($\bar{e}_{i,t+1}$) is a probability threshold that allows to distinguish between those who face an above average risk of being poor next year and those who face a below average risk of falling into poverty (the more secure). However, as I explain below, my objective is to derive a vulnerability threshold expressed in terms of income. Therefore, I derive the vulnerability line by calculating the incomes associates with the relevant poverty entry risks.

My approach, like that of López-Calva & Ortiz-Juárez (2014), Dang & Lanjouw (2017) and Schotte et al. (2018), has an implicit monotonicity assumption: the higher the income -above the poverty line- the lower the poverty entry probability. Although this assumption is plausible, it does not guarantee that higher base period income (among the non-poor) implies a lower probability of falling into poverty.

There are two reasons to propose a vulnerability threshold in terms of household income even though a monotonic relationship assumption may not always apply. First, using vulnerability line instead of the probability threshold estimated in the first step facilitates its interpretation for social protection and poverty reduction policies because it has a natural compatibility with the poverty line used in its calculation (Dang & Lanjouw, 2017). Second, as it happens with poverty measures, where theory supports the selection of the poverty line cut-off criterion (e.g. basic needs approach), the vulnerability line measures connect with the well-defined notion of vulnerability to poverty approach (López-Calva & Ortiz-Juarez, 2014), which sets a criterion to estimate the lower-threshold of the middle class. By doing so, my measures deal with the economic literature that uses income thresholds to define the middle class (e.g. Banerjee & Duflo (2008); Birdsall (2010)).

Furthermore, following López-Calva & Ortiz-Juarez's (2014) argument, I believe it is important to use predicted income rather than observed average income because the outcome of a parameterised model is less volatile than the observed values.[16] Therefore, I can assume that a predicted household income better reflects the household income generation capacity because it is related to its composition, the types of assets owned by the household, and its environment (location of the house).

I calculate a vulnerability line for a non-poor sample as follows:

I use a log-linear regression model to estimate a cross-sectional household income equation for the

estimates because a large number of observations is dropped, thereby making the sample less representative (Kanabar, 2017).

[16] One weakness of the Schotte et al. (2018) approach is the use of observed household expenditure to estimate the vulnerability line.

base year at the household level. I use the same time-fixed predictor variables as in the endogenous switching model in the following expression:

$$ln\ y_{i,t} = \beta X_{i,t} + \varepsilon_{i,t} \quad (8)$$

where $\ln y_{i,t}$ is the log of household per capita income for year t. I predict household per capita income for year t, for each non-poor household i, based on the coefficient estimates from eq. (8).

I predict $\ln y_{i,t}$ taking into account the retransformation problem of $\ln y_{i,t}$ (Santos Silva & Tenreyro, 2006). I tackle the problem by applying Duan's (1983) solution.[17] I fit the log-linear regression using Poison regressions methods as a way of obtaining estimates of $y_{i,t}$, namely:

$$y_{i,t} = \exp\ (\beta X_{i,t} + \varepsilon_{i,t}) \quad (9)$$

That is, instead of taking the expectation of $\ln\ y_{i,t}$, I estimate the expected value of $y_{i,t}$.

$$E(y_{i,t}) = \exp\ (\beta X_{i,t})E\{\exp(\varepsilon_{i,t})\} \quad (10)$$

Assuming that $\varepsilon_{i,t}$ is independent and identically distributed, I estimate $E\{\exp\ (\varepsilon_{i,t})\}$ by the sample average $\mathrm{N}^{-1}\sum_{i=1}^{N}\exp\ (\widehat{\varepsilon_{i,t}})$.[18]

At the final, third step, I calculate the vulnerability line (V_t) as the mean predicted per capita income at t for non-poor households (k) with a predicted households' probability to enter into poverty in t that falls $\pm$ 1 percentage points probability around the poverty entry rate $(\bar{e}_{t+1} \pm 0.01)$. That is:

$$V_t = E(y_{k,t}|\mathrm{x}_{k,t}) = \exp\ (\beta \mathrm{x}_{k,t})E\{\exp(\varepsilon_{k,t})\} \quad (11)$$

for all $e_{k,t+1}|\mathrm{x}_{k,t}$ (as defined in equation (7)) such that:

$$\bar{e}_{t+1} - 0.01 \leq e_{k,t+1}|\mathrm{x}_{k,t} \leq \bar{e}_{t+1} + 0.01\ .$$

The vulnerability line (V_t) obtained using a range around the poverty entry rate allows to reduce the volatility of the risk cut-off point and to provide enough observations to get a robust estimate

[17] López-Calva & Ortiz-Juarez (2014) model for predicting household income neglects the retransformation problem. They obtain the vulnerability line $(E(y_{i,t}|X_{i,t}))$ assuming a straightforward retransforming of the income scale. However, they predict $E(ln\ y_{i,t}|X_{i,t})$ and take the exponent as result, which is incorrect because the expected value of the logarithm of the variable of interest is different from the logarithm of its expected value $(E(y_{i,t}|X_{i,t}) \neq \exp\{E(ln\ y_{i,t}|X_{i,t})\})$. Thus, it biases estimates of household income.

[18] When comparing the average of the observed income of the household using the base year t of the survey Panel CASEN with a simple prediction, we obtain a difference of 15.2 per cent between the two values (CL $158,215, and CL $134,126, respectively). When using Duan's (1983) method to address the retransformation problem, the prediction of the average is CL $159,300. This value differs by less than 0.01 per cent from the sample mean value, thus showing that ignoring the retransformation bias leads to a poor prediction of household income.

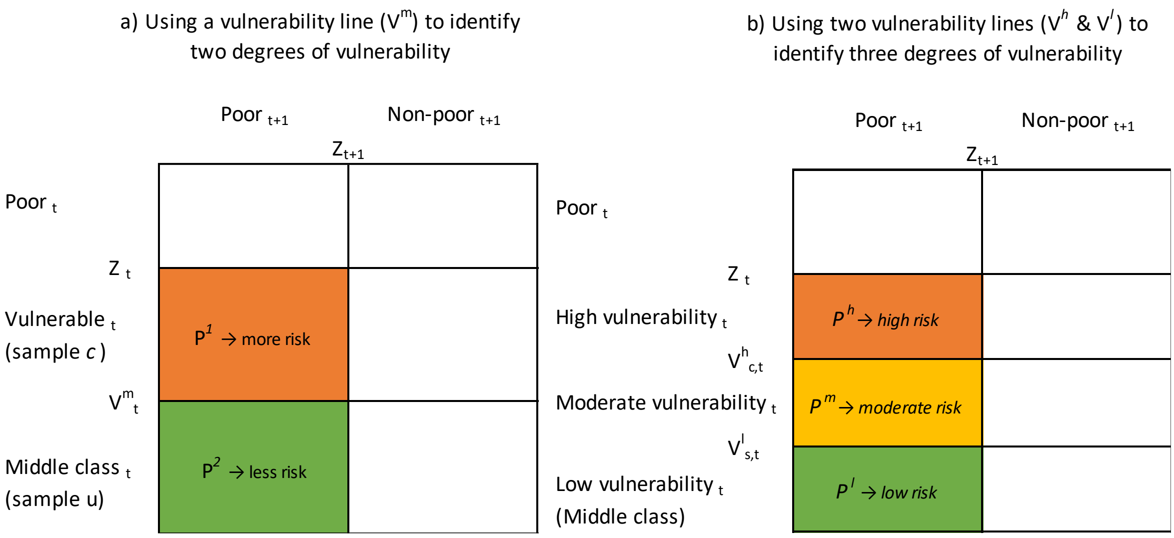
of $E(Y_{k,t}|x_{k,t})$ in Eq. 11. This strategy is both independent of the size and design of the panel sample and it provides similar vulnerability thresholds when I use narrower or wider percentage points probability bands.[19] It is worth mentioning that vulnerability lines are sensitive to the household income used in their calculation (observed income, predicted income, and predicted income addressing the retransformation bias). See Table 2 in section 5 for both sensitive analyses.

Extension: Using more than one vulnerability line to classify social groups according to their degrees of vulnerability to poverty

Although my approach allows for the calculation of several vulnerability lines, I derive only two, as this is the minimum number of cut-off points to keep the number of income groups identified manageable for policy purposes. The two vulnerability lines allow to distinguish within the non-poor population groups three levels of risk of falling into poverty: high, moderate and low.

One vulnerability line focuses on households that are located in the central part of the income distribution (sample c), and the other line focuses on the upper part of the income distribution (sample u). The mobility matrix in Figure 1. describes how I identify three degrees of vulnerability to poverty using two vulnerability lines.

Figure 1: Mobility matrices to illustrate how to identify degrees of vulnerability to poverty



First, I calculate a moderate (m) vulnerability line ($V^m_{n,t}$) for all non-poor households in the base year (sample n, which does not include the rich) using Eq. 7. This moderate vulnerability line is

[19] Another weakness of the Schotte et al. (2018) approach is that the use of the confidence interval to estimate the vulnerability line is undesirable, as the poverty risk range is a function of sample size and design. For example, it is unclear how to obtain the vulnerability line when no sample observation falls within the estimated confidence interval.

associated with the poverty entry rate ($\bar{e}_{n,t+1}$) and it allows to split sample n in two sub-samples in time t: i) sample c with households with their income between the poverty line (Z_t) and $V^m_{n,t}$; and ii) sample u with households with their income above $V^m_{n,t}$. Assuming that increase in income can lower the probability of poverty entry, the probability of falling into poverty for all households in sample C is higher than $\bar{e}_{n,t+1}$, and for all households in sample u is lower than $\bar{e}_{n,t+1}$.

Including the moderate vulnerability line allows to estimate two vulnerability indexes or transition proportions shown in the mobility matrix of Panel A in Figure 1. One is the vulnerable index (P^1) and the other is the insecurity index (P^2). P^1 and P^2 correspond to the expected proportions of those falling into poverty at more and less risk than the average, orange cell and green cell, respectively.

$$P^1 = P(y_{t+1} \leq Z_{t+1} \,|\, Z_t < y_t \leq V^m_{n,t}) \tag{12}$$

$$P^2 = P(y_{t+1} \leq Z_{t+1} \,|\, V^m_{n,t} < y_t) \tag{13}$$

P^1 and P^2 are transition proportions in a mobility matrix similar to Dang & Lanjouw's (2017) vulnerability indexes.[20] However, unlike their approach in which the proportions in the matrix are given, and the vulnerability line is derived, I estimate both vulnerability indexes from my moderate vulnerability line ($V^m_{n,t}$).

Second, since my approach (step 1 and 2) makes it possible to obtain vulnerability lines for different non-poor populations in the base year, I can simultaneously obtain a high vulnerability line ($V^h_{c,t}$) associated with the poverty entry rate ($\bar{e}_{c,t+1}$) for households in the central part of the income distribution (sample C), and a low vulnerability line ($V^l_{u,t}$) associated with the poverty entry rate ($\bar{e}_{u,t+1}$) for those in the upper part of the income distribution (sample u).

The mobility matrix of Panel B in Figure 1 shows how the high vulnerability line and low vulnerability line allow to estimate three vulnerability indexes: the high vulnerability index P^h (orange cell); the moderate vulnerability index P^m (yellow cell); and the low vulnerability index P^l (green cell).

P^h corresponds to the expected proportion of falling into poverty in $t+1$ of those at high risk ($\bar{e}_{c,t+1} \leq e_{i,t+1}$, assuming a monotonic relationship between poverty risk predicted and income).

---

[20] Dang & Lanjouw (2017) provide two measures of vulnerability to poverty: the "insecurity index" and "vulnerability index", "but the insecurity index focuses on households in the top part of the consumption distribution while the vulnerability index focuses instead on those located in the middle" (Dang & Lanjouw, 2017, p. 639). These authors approach offers greater flexibility in defining vulnerability to poverty, yet, in practice, they use a single income threshold.

$$P^h = P(y_{t+1} \leq Z_{t+1} | Z_t < y_t \leq V^h_{c,t}) \tag{14}$$

P^m is the transition probability for non-poor people with a moderate risk of falling into poverty ($\bar{e}_{u,t+1} \leq e_{i,t+1} < \bar{e}_{c,t+1}$, ditto).

$$P^m = P(y_{t+1} \leq Z_{t+1} \; | V^h_{c,t} < y_t \leq V^l_{u,t}) \tag{15}$$

Finally, P^l corresponds to the expected probability of being poor in $t+1$ for those with a low risk ($e_{i,t+1} < \bar{e}_{u,t+1}$, ditto).

$$P^l = P(y_{t+1} \leq Z_{t+1} \; | V^l_{u,t} < y_t) \tag{16}$$

My approach can be easily adapted to derive more than two vulnerability lines (e.g. for income quintile or decile groups) in case it would be required in a particular application.[21]

4. The case of Chile: data and definitions

I apply the framework described above to Chile. This country shows some specific characteristics that makes it a compelling case to derive the vulnerability lines. First, in 2013 Chile was classified by the World Bank as a high-income country, reaching a Gross National Income per capita of around US$13,000 adjusted by international inflation (Tezanos & Sumner, 2016). As a consequence of this economic progress and its highly focused social policies, Chile has experienced a remarkable decline in poverty over the last decades (Cingano, 2014; Larrañaga & Rodríguez, 2015).[22] However, several studies reported that the improvement of this measure of economic well-being was accompanied by a generalised social discontent with the economic and political model (e.g. PNUD (2017)). This was evidenced by the massive protests that started in October 2019 when an increase in the public transport fare was announced (Pons, Mullins, Masko, Lobb, & Tella, 2020).

Second, the progress of the Chilean society towards higher levels of social inclusion has been limited. Based on post-transfer and post-tax household income per capita, official data from Chile show that the Gini coefficient decreased only two points between 1990 and 2017, from 0.521 to 0.502 (MDS, 2018). These figures are among the highest among OECD countries (OECD, 2018b).

[21] This feature might suggest that if, in the limit, I end up using the poverty risk (or corresponding income) information as a continuous, my approach would not be different from a VEP approach. However, this is not the case. In the VEP approach the probability of a household being poor refers to all current poor and non-poor households in its estimation. In my approach, the relevant population is the currently non-poor households.

[22] According to the official poverty measure used by the Chilean government during this period, the share of people living below the national absolute poverty line decreased from 38.6 per cent in 1990 to 8.6 per cent in 2017 (MDS, 2018).

The high level of inequality reflects a large gap between the top and mean incomes (Chauvel, 2018). As a result of this gap, the income distribution is narrower in the lowest decile groups with a high turnover of many households around the absolute poverty line (Denis, Prieto, & Zubizarreta, 2007; Larrañaga, 2009). This characteristic of the Chilean income distribution suggests that many households are extremely vulnerable to falling into poverty (Maldonado, Prieto, & Lay, 2016; Neilson, Contreras, Cooper, & Hermann, 2008).

Third, Chile conducted a household panel survey between 2006-2009. It is a household survey that collected data each year over a period of four years, providing a great opportunity to study the dynamics of poverty in Chile in order to propose vulnerability lines to study both the vulnerable group and the middle class.

Data and definition of income poverty

For the analysis presented in this paper, I exploit the rich data set of the Chilean Socioeconomic Household Panel Survey (P-CASEN) for the years 2006, 2007, 2008 and 2009.[23] The P-CASEN is a household-based panel study that collected information related to income, education, employment, health, household composition, and housing (Observatorio Social, 2011c). The interviews were conducted annually with all members of each household (adults and children). The target population consisted of all private households throughout the national territory. For the selection of cases, the National Socioeconomic Characterization Survey (CASEN) 2006 was used as the sampling frame. The first round of the P-CASEN in 2006 consisted of 8,079 households, comprising a total of 30,104 individuals (Lynn, Zubizarreta, & Castillo, 2007). The main advantage of this dataset is that it follows individuals and households over time.

Although the P-CASEN is a longitudinal survey at the individual level, the explanatory variables that are part of the poverty entry equation are measured at the household level (equation 3). This choice is necessary because subsequent analyses require the same variables at the household level (equation 11). Following the strategy of Schotte et al. (2018), we checked the robustness of this approach with various alternative specifications, such as individual-level controls or considering only a sample with household heads. The parameter estimates are consistent across all specifications, suggesting that there is no systematic bias in the estimated coefficients.[24]

A relevant methodological decision is whether or not to work with a sample restricted to the adult

[23] The datasets analysed in the study were taken from the following public domain resources: http://observatorio.ministeriodesarrollosocial.gob.cl/encuesta-panel-casen-2006. More information on the CASEN panel can be found at: http://observatorio.ministeriodesarrollosocial.gob.cl/enc_panel.php

[24] The regression results of the robustness tests are available from the author if requested.

population. In most studies of the dynamics of poverty, the analysis is limited to the population aged between 25 and 64 years (Ayllón, 2013; Buddelmeyer & Verick, 2008; Cappellari & Jenkins, 2004). The justification for this is that children and young people under 26 do not have an impact on decisions related to the income of the household. Also, by not including individuals over 64 years of age, researchers aim to avoid the impact of retirement on poverty dynamics transitions, particularly the impacts of pensions on income levels. Yet, the studies that propose vulnerability lines generally do not limit the age of adults for their analysis (Dang & Lanjouw, 2017, 2017; López-Calva & Ortiz-Juarez, 2014; Zizzamia, Schotte, Leibbrandt, & Vimal Ranchhod, 2016). Therefore, given that one of the objectives of this research is to compare the results obtained with these types of works, I consider all of the adult population.

In this research, the welfare of individuals is named in terms of monthly income. Specifically, the income corresponds to the sum of the income of the household (mainly salaries, wages and earnings from independent work), cash transfers received from social programmes, and the imputation of the rent when the house is inhabited by its owners. November was the reference month for questions about net income (after taxes). Questions without answers and values lost in the components that form the income have been solved by using imputation procedures (Observatorio Social, 2011b).

To identify the low-income population, I use two absolute poverty lines. This procedure implies identifying the poor using the same income cut-off for each round. The first absolute cut-off is the official line of urban poverty in Chile in 2009, which in Chilean pesos (CL$) corresponds to a monthly income of CL$ 64,134 ($6.41 dollars per person per day (pppd) in 2011 purchasing power parity (PPP)). This poverty line was defined according to the minimum monthly income established per person to satisfy basic needs, which was calculated by ECLAC (Mideplan, 2010).

The second income cut-off corresponds to the international poverty line recommended by the World Bank to compare levels of poverty in countries in Latin America that are considered upper-middle income. Even though Chile is a high-income country according to the World Bank, I do not use the poverty line for this group of countries because it is too high to be applied to Chile. Instead, I use the poverty line for upper-middle-income countries, which better fits the Chilean context. This value is $5.5 dollars pppd in 2011 PPP terms. This threshold is based on the work of Jolliffe & Prydz (2016), who linked the poverty lines of 115 countries that are close to the 2011 PPP reference period with the income levels of each country, proposing four international poverty lines for four country categories: low income, lower-middle income, upper-middle income, and high income.

5. Modelling poverty entries: Testing specifications and results of the estimates

In this section, I present the results of the endogenous switching model. First, I test the specifications of the model that will allow to estimate the probability of falling into poverty for non-poor households. For that, I compare the estimates of the model with the data, estimate the correlation between unobservables, and perform several tests to determine the ignorability of both initial conditions and attrition. Second, I present the results of the estimates for the conditional poverty equation using the poverty line recommended by the World Bank for upper middle-income countries.

Testing model specification

First, I present an assessment of the degree of fit of the model to the CASEN data panel. Panel 1 in Table 1 presents the predictions that the model calculates from equation (3) for the official poverty line in Chile. The overall average of individuals that enter poverty in period $t + 1$ (since they were not poor in period t) is 0.146, which is close to the 0.142 from the matrix of annual poverty transitions. For the proportion of individuals that remain in the panel sample, the value of the predicted probability and the raw value are both 0.834. The same is true for the initial poverty ratio (0.257). These predictions show that the specified model replicates the sample averages closely.

Table 1: Predicted probabilities, estimates of the model correlations and statistics tests

1. Predicted probabilities	Estimate		Std. Dev.
Poverty entry	0.146		(0.106)
Initially poor	0.257		(0.206)
Survey retention	0.834		(0.183)
2. Correlations between unobservable components			
ρ_1: Initial and conditional poverty	0.043		(0.044)
ρ_2: Survey retention and initial poverty	0.025	**	(0.012)
ρ_3: Survey retention and conditional poverty	0.032	**	(0.013)
3. Wald test of correlations (null hypotheses for tests)	Test statistic		p-value
$\rho_1 = \rho_2 = 0$: No evidence of initial conditions	6.48	**	0.0391
$\rho_1 = \rho_3 = 0$: No evidence of non-random attrition	10.12	***	0.0064
$\rho_1 = \rho_2 = \rho_3 = 0$: Joint exogeneity	10.71	**	0.0134

Source: Author's calculations based on the P-CASEN 2006-2009.

Notes: Asymptotic standard errors are robust to repeated observing the same individual. Simulated pseudo maximum likelihood estimation with 250 random draws. *** significance at 1 percent; ** significance at 5 percent; * significance at 10 percent.

One of the advantages of using a first-order Markov approach is that it takes into account the initial conditions and non-random survey attrition. In order to evaluate the possible ignorability of these two selection mechanisms in the model, I test for the separate and joint significance of the correlation coefficients associated with the selections in equations (1) and (2). The term ignorability here means that the different equations of the model can be estimated separately without worrying that the estimates are biased.

As illustrated in Panel 2 in Table 1, there is no significant evidence of an unobserved correlation ρ_1 between initial and conditional poverty in the P-CASEN data. However, there is strong statistical evidence that the unobservable factors of non-random attrition are positively correlated with both the initial poverty in the base year ρ_2 and with the conditional poverty status ρ_3.

These results should not be surprising because they confirm a greater retention in the panel sample of those who were poor initially compared with those who were non-poor and also those who were poor in the next period compared to those who were above the poverty line. This result implies that the sample panel contains a non-random attrition problem. The exogeneity tests of the two selection processes considered could be rejected by the Wald tests conducted. Thus, both initial condition of poverty status and survey retention could be regarded as endogenous to the model (see panel 3 in Table 1).

In summary, the tests in the correlations of the unobservable factors indicate that the initial condition and the attrition of the sample are endogenous. Therefore, it is necessary to use the three equations (1, 2 and 3) of the endogenous switching framework to estimate the entry rates into poverty.

The drivers of poverty entry

Table 2 shows the coefficients for the probability of entering poverty from equation (7) using the poverty line of $5.5 dollars pppd in 2011 PPP terms, which corresponds to the cut-off suggested by the World Bank to compare upper-middle income countries.[25]

In terms of the characteristics of the household head, those who are less likely to fall into poverty are older males with a university education. Work wise, for heads of households in both informal jobs and for the unemployed the conditional probability of poverty entry is higher.

The characteristics of the partner of the household head that affect poverty entry differ in some respects from the characteristics of the heads of households. In this regard, when the partner has

[25] In the appendix section, Table B1 shows the coefficients using the official line of urban poverty in Chile.

a university degree has a greater impact on reducing the risk of the household of entering into poverty than when the head of household has a university degree. When the partner is inactive it has a significant impact on increasing the household's likelihood of entering poverty. On the contrary, although working in an informal job and unemployment are both statistically significant, they have a lower weight in explaining falls in poverty than in the case of the head of the household. These results confirm the findings found in other studies on poverty dynamics carried out in Chile (e.g. Denis et al., 2007; Maldonado et al., 2016).

Finally, regarding the characteristics of the household, singles with children have a higher risk of falling into poverty. The same counts for larger families with more children. The attributes that reduce the risk of falling into poverty are: (i) the number of working household members, (ii) owning the house where they live (or paying a mortgage) and, in terms of location, (iii) living in urban areas and regions 11 and 12. Similar results are found in the works of Neilson et al. (2008) and Maldonado & Prieto (2015).

As I have already explained, the model controls for the endogeneity of poverty status in the initial period (equation 1) and non-random attrition (equation 2). When looking at column of poverty entry in Table 2, it can be seen that most of the covariates that are statistically significant in the association with initial poverty in time (t) are also significant in the case of conditional poverty status. It should be mentioned that covariates, such as having a university degree or the number of individuals working in the household, have a larger impact on the increase and decrease in the risk of being poor in the base year, than in the case of the equation to estimate the chances of falling into poverty.

As to the exclusion restrictions used in this equation, it stands out that when the mother of the individual surveyed works as a salaried employee this increases the probability of being poor in the base year, whereas when the father also works as a salaried employee the likelihood of entering poverty decreases. In the case of the education levels of the parents of the interviewee, both parents have a negative impact on the initial condition of being poor when the parents have only finished secondary education.

Column of survey retention of Table 2 shows the factors that explain the attrition of the P-CASEN sample. The two characteristics of the heads of households that make them less likely to be retained in the sample in the following period are (i) being male and (ii) having a university degree. The occupational categories do not seem to have a significant impact on attrition. In the case of the characteristics of the partner of the household head that increase the probability of remaining in the sample, these are (i) having completed only primary education, and (ii) being unemployed or inactive.

In the case of the characteristics of the household, being a household that is single with children has a positive impact on retention. Conversely, for single-person households, the impact is negative. Lastly, the variable that indicates whether the individual is an original member of the sample has a positive and the highest coefficient, which indicates that an individual who interviewed in the first round has a high probability of not leaving the sample in the next period.

Table 2: Model estimates of poverty entry rates, initial poverty status and survey retention, Chile (2006-2009)

Variables (measured at *t*)	Poverty entry: Poor at t+1 | Non-poor at t			Poverty status at t			Survey retention		
	Coefficient		Std. Dev.	Coefficient		Std. Dev.	Coefficient		Std. Dev.
Household head characteristics									
Female	0.039	*	(0.020)	0.171	***	(0.020)	-0.046	**	(0.021)
Age	-0.006	***	(0.001)	-0.014	***	(0.001)	0.001		(0.001)
Education: Ref. Secondary school									
Primary school	0.140	***	(0.020)	0.362	***	(0.018)	0.133	***	(0.020)
University degree	-0.331	***	(0.037)	-0.679	***	(0.047)	-0.163	***	(0.029)
Labour status: Ref. Formal employed									
Informal employed	0.416	***	(0.025)	0.555	***	(0.022)	0.011		(0.025)
Unemployed	0.425	***	(0.093)	0.993	***	(0.055)	-0.004		(0.076)
Inactive	-0.048		(0.041)	0.146	***	(0.033)	-0.006		(0.037)
HH head's partner characteristics									
Age	-0.008	***	(0.001)	-0.005	***	(0.001)	-0.001		(0.001)
Education: Ref. Secondary school									
Primary school	0.240	***	(0.023)	0.329	***	(0.023)	0.045	*	(0.024)
University degree	-0.400	***	(0.063)	-0.742	***	(0.121)	-0.056		(0.042)
Labour status: Ref. Formal employed									
Informal employed	0.255	***	(0.033)	0.376	***	(0.039)	-0.004		(0.035)
Unemployed	0.168	***	(0.052)	0.350	***	(0.043)	0.227	***	(0.057)
Inactive	0.100	***	(0.022)	0.184	***	(0.022)	0.046	**	(0.022)
Household characteristics									
Household type: Ref. Couple without children									
Single without children	0.153	***	(0.034)	0.121	***	(0.038)	-0.033		(0.033)
Couple with children	0.136	***	(0.029)	0.255	***	(0.030)	0.030		(0.027)
Single with children	0.329	***	(0.035)	0.546	***	(0.036)	0.155	***	(0.035)
Lone person	0.021		(0.062)	-0.083		(0.065)	-0.198	***	(0.056)
Number of persons	0.054	***	(0.008)	0.274	***	(0.007)	-0.028	***	(0.007)
Number of children < 15	0.140	***	(0.013)	0.085	***	(0.012)	0.065	***	(0.012)
Number of workers	-0.201	***	(0.017)	-0.930	***	(0.015)	0.009		(0.010)
Housing: Ref. Own housing (mortgage)									
Own housing, mortgage	-0.366	***	(0.030)	-0.415	***	(0.034)	-0.149	***	(0.025)
Rent	0.217	***	(0.026)	0.380	***	(0.026)	-0.364	***	(0.024)
Subsidized or rent free	0.204	***	(0.025)	0.661	***	(0.019)	-0.075	***	(0.023)
Rural	0.133	***	(0.023)	0.154	***	(0.022)	0.094	***	(0.026)
Regions: Ref. 13th									
1st, 2nd, 3rd and 4th	0.094	***	(0.026)	0.083	***	(0.026)	0.034		(0.025)
5th, 6th, 7th, 8th, 9th and 10th	0.152	***	(0.018)	0.275	***	(0.019)	0.147	***	(0.018)
11th and 12th	-0.215	***	(0.050)	-0.287	***	(0.052)	0.271	***	(0.055)
Time (*t*): Ref. 2007									

2008	0.105	***	(0.018)	-0.173	***	(0.016)			
2009	-0.154	***	(0.019)	-0.007		(0.016)			
Individual characteristics (Exclusion restrictions)									
Mother education: Ref. No schooling									
Primary school				-0.050		(0.033)			
Secondary school				-0.156	***	(0.047)			
University degree				-0.155		(0.103)			
Type of work done by mother: Ref. Self-employed									
Employership				0.005		(0.123)			
Paid employment				0.116	***	(0.033)			
Non-employment				0.018		(0.028)			
Father education: Ref. No schooling									
Primary school				0.016		(0.034)			
Secondary school				-0.107	**	(0.045)			
University degree				0.028		(0.098)			
Type of work done by father: Ref. Self-employed									
Employership				-0.115		(0.073)			
Paid employment				-0.069	***	(0.026)			
Non-employment				0.014		(0.108)			
Original sample member							0.509	***	(0.056)
Constant				-1.181	***	(0.063)	0.602	***	(0.081)
Log-pseudolikelihood				-61,078.240					
Wald chi-square (d.f. = 131)				316,449.326 (p<0.000)					
Number of observations (person-waves)				65,205					

Source: Author's calculations using the P-CASEN 2006-2009.

Notes: Asymptotic standard errors are robust to repeated observing the same individual. Simulated pseudo maximum likelihood estimation with 250 random draws.

*** significance at 1 percent; ** significance at 5 percent; * significance at 1 percent.

5. Predicting vulnerability lines using the low-income dynamic model estimates

In this section, I present the results of the poverty dynamic approach to identifying the degrees of vulnerability to poverty in the distribution of income in Chile. The discussion of these results is presented in two sub-sections. First, I show the vulnerability lines associated with the risk of falling into poverty for each sample of households specified in the three stages of my proposal. Second, I use the low and high vulnerability lines to classify currently non-poor people into three risk groups: low, moderate and high risk of falling into poverty in the next period.

Vulnerability lines by poverty entry rates

In this sub-section, I present the results of predicted household income by poverty entry rates for different non-poor samples. Table 3 shows vulnerability lines in the base year for three subsamples of non-poor associated with the average probability of falling into poverty next year. When using the World Bank poverty line ($5.5 dollars pppd in 2011 PPP), the moderate vulnerability line is

$12.8 dollars pppd with a poverty entry rate of 11.2 per cent for the all non-poor. The value obtained do not differ much from the $13.0 pppd delivered by the World Bank (2018) after its most recent update of the vulnerability line. However, the interpretation offered by the World Bank differs significantly from the one obtained from my result. While the vulnerability line of the World Bank is associated with a risk of falling into poverty of 10 percent in a time horizon of between 3 and 5 years (López-Calva & Ortiz-Juarez, 2014), I obtain a vulnerability line related to the average risk that households have of falling into poverty from one year to the next.

Table 3: Vulnerability lines for subsamples of non-poor in the base year (t) using different poverty lines

Vulnerability lines for different poverty lines	Sub-samples of non-poor in the base year (t)	Poverty entry rate next year (t+1)				Vulnerability line in base year (t)			
		Mean	Std. Dev.	[95% Conf.	Interval]	Mean	Std. Dev.	[95% Conf.	Interval]
Z_t: \$5.5 pppd in 2011 PPP									
1. Moderate vulnerability line (V_t^m)	$Z_t < y_t$	0.112	0.001	0.109	0.115	12.77	0.11	12.54	12.99
2. Low vulnerability line	$V_t^m < y_t$	0.046	0.001	0.043	0.049	20.03	0.17	19.71	20.36
3. High vulnerability line	$Z_t < y_t < V_t^m$	0.171	0.002	0.167	0.176	9.86	0.06	9.75	9.97
Z_t: \$6.41 pppd in 2011 PPP									
1. Moderate vulnerability line (V_t^m)	$Z_t < y_t$	0.138	0.002	0.135	0.141	11.79	0.05	11.69	11.90
2. Low vulnerability line	$V_t^m < y_t$	0.067	0.002	0.064	0.070	17.40	0.07	17.26	17.55
3. High vulnerability line	$Z_t < y_t < V_t^m$	0.233	0.003	0.223	0.239	8.57	0.06	8.45	8.69

Source: Author's calculations using the P-CASEN 2006-2009.

Note: To describe the sub-samples I used the following notation: Z_t (poverty line in year *t*) y_t (household income in year *t*), V_t^m (moderate vulnerability line in year *t*).

The low vulnerability line enables to identify the income-secure middle class. The second line in Table 3 shows that the income threshold for the lower bound for this group is $20.0 dollars pppd with an average probability of falling into poverty of 4.6 per cent. This value is a third higher than the vulnerability line used by the World Bank for the same purpose, namely, to be the lower limit to identify those who are the income-secure middle class due to having a low risk of falling into poverty.

Furthermore, the low vulnerability line that I propose is close to the $21.19 dollars pppd in 2011 PPP terms of the poverty line used to compare high-income countries (Jolliffe & Prydz, 2016). In this way, the cut-off line to define the income-secure class in upper-middle-income countries would provide a direct association with the absolute poverty line of high-income countries that could be used in future research to study and compare changes in the income distribution among high-income countries with upper-middle-income countries.

The high vulnerability line is 9.9 dollars pppd and the poverty entry rate for the non-poor sub-sample is 17.1 per cent. The fact that the value of the high vulnerability line is a 30 per cent lower than the vulnerability line for the non-poor should not be a surprise. As discussed below, this is due to the high proportion of the non-poor population that is very close to the poverty line.

For the Chilean official poverty line ($6.41 dollars pppd in 2011 PPP) the moderate vulnerability line is $11.8 dollars pppd, the low vulnerability line is $17.4 dollars pppd, and the high vulnerability line is $8.6 dollars pppd. As expected, these income cut-offs are lower than the vulnerability line based on World Bank poverty line used in upper-middle-income countries.

Sensitivity analysis of the vulnerability line approach to associate poverty entry rates with household income

I assess the sensitivity of the calculated vulnerability lines to some of the choices I made in deriving them. First, I evaluate the sensitivity to the selection of using the ± 1 per cent interval to calculate the average monetary threshold associated with a poverty entry rate. Panel A in Table 4 shows that a choice of a narrower probability interval of ± 0.5 per cent would have led to similar income cut-offs for both high vulnerability line and low vulnerability line. For a wider interval around the poverty entry rate such as ± 2 per cent the high vulnerability line and low vulnerability line change less than 3 per cent compared with the vulnerability lines for the ± 1 per cent interval.

Table 4: Sensitive analysis for the association of vulnerability lines with predicted poverty entry rates

Sensitive analysis for vulnerability lines	High vulnerability line associated with a poverty entry rate of 4.6 per cent				Low vulnerability line associated with a poverty entry rate of 13.4 per cent			
	Mean	Std. Dev.	[95% Conf. Interval]		Mean	Std. Dev.	[95% Conf. Interval]	
a) Percentage points probability around the poverty entry rate								
± 0.5	9.88	0.07	9.74	10.03	20.15	0.12	19.91	20.40
± 1	9.86	0.06	9.75	9.97	20.28	0.09	20.10	20.46
± 2	9.92	0.04	9.84	9.99	20.85	0.07	20.71	20.98
b) Household income								
Observed income	8.30	0.05	8.21	8.40	23.17	0.17	22.84	23.49
Predicted income without addressing the retransformation bias	8.38	0.05	8.29	8.47	17.23	0.08	17.08	17.39
Predicted income addressing the retransformation bias	9.86	0.06	9.75	9.97	20.28	0.09	20.10	20.46

Source: Author's calculations using the P-CASEN 2006-2009.

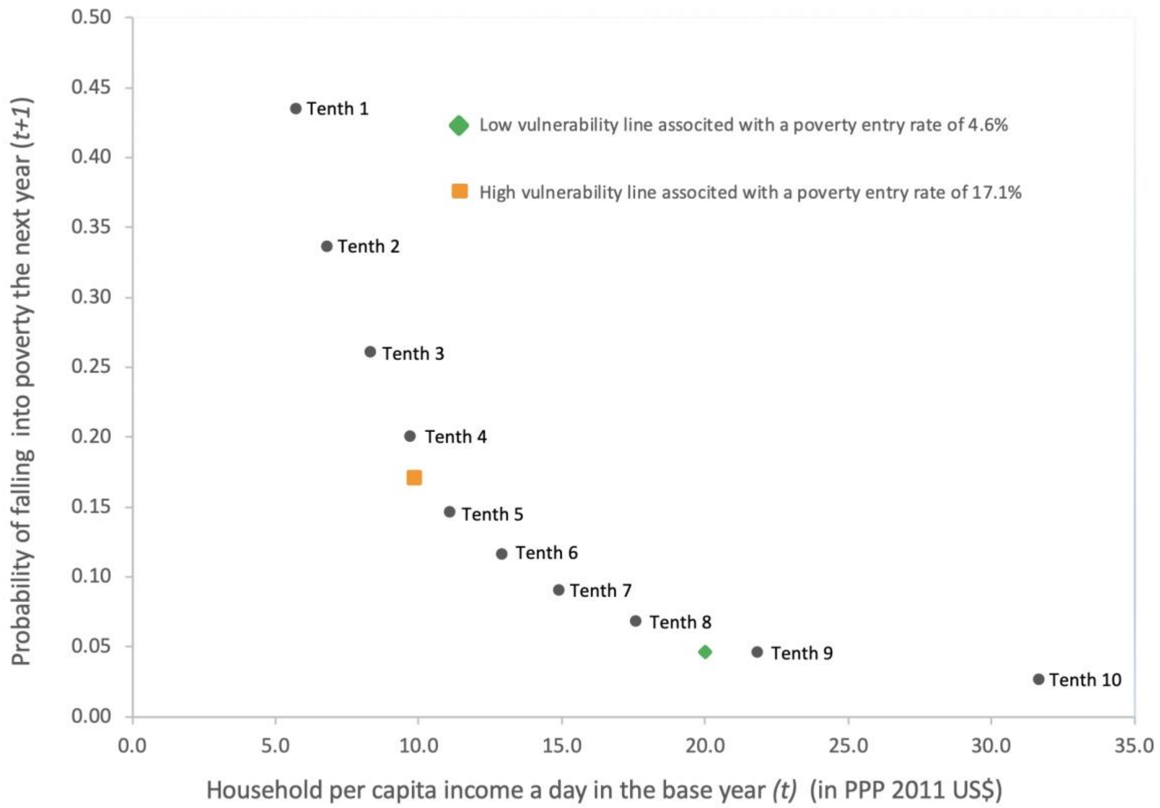
Notes: Vulnerability lines derived from the World Bank poverty line ($5.5 dollars pppd in 2011 PPP). In Panel B, all household income used the ± 1 per cent interval around the poverty entry rate.

Second, I assess the difference between vulnerability lines depending on the household income used (see panel B, Table 4). My high vulnerability line is around 17 per cent higher than the high vulnerability lines that are calculated using the observed income and the predicted income (without addressing the retransformation bias). When I compare the low vulnerability lines, the differences are even more significant. My low vulnerability line is 34 per cent higher than the low vulnerability line based on the predicted income without addressing the retransformation bias and 18 lower than the one calculated with observed income.

Using both high and low vulnerability lines to measure high, moderate and low vulnerable

Based on the results of Table 1 and using the World Bank poverty line, Figure 2 shows the poverty entry rate associated with both the low and high vulnerability lines. It also shows the association between the average risk of falling into poverty from one year to the next and the level of household income for each of the tenths of the income distribution. Although the deciles correspond to the entire income distribution, the subsample of decile 1 considers only those that were non-poor in the period t, and decile 10 considers only those households that had an income inferior to $70 pppd.

Figure 2: Vulnerability lines by poverty entry rates for non-poor subsamples



Source: Author's calculations using the P-CASEN 2006-2009.

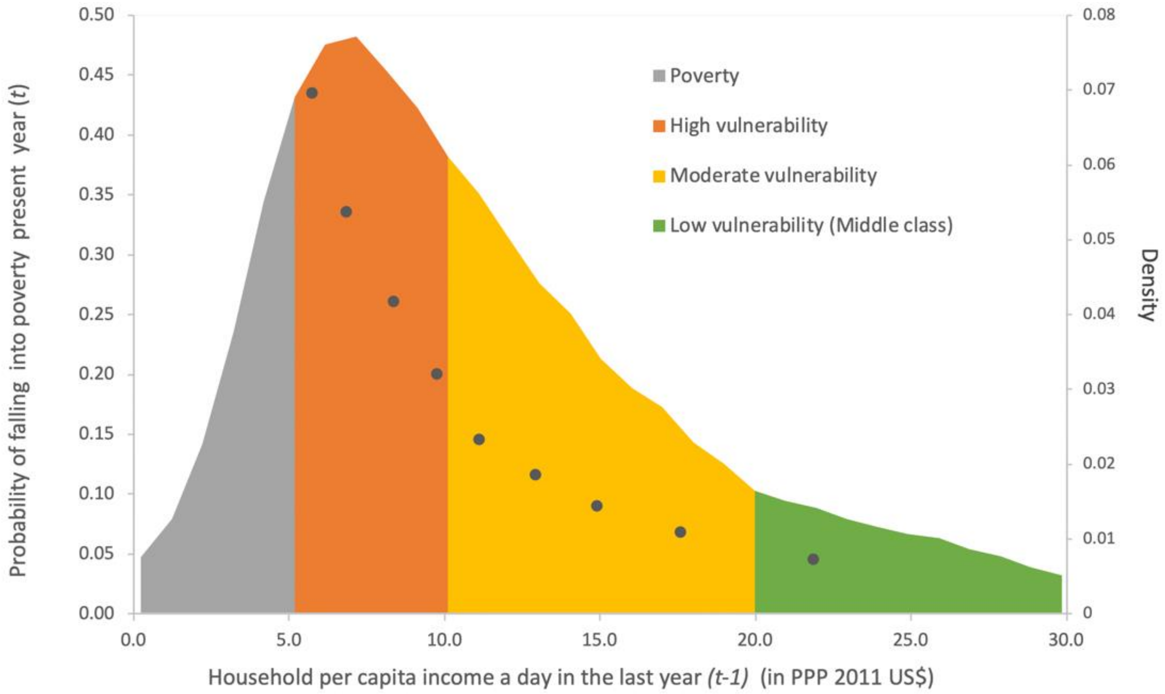
Note: I used the World Bank upper middle-income countries poverty line ($5.5 dollars pppd in 2011 PPP).

The green diamond in Figure 2 shows the low vulnerability line ($20.0 dollars pppd) associated with its probability of entering into poverty (4.6 per cent). The absolute cut-off for the low vulnerability line is in between the average risk of falling into poverty of decile groups 8 and 9 of the income distribution. This indicates that less than 20 percent of the population in Chile can be considered part of an income-secure middle class.

The orange square indicates the high vulnerability line ($9.9 dollars pppd) associated with its probability of entering into poverty (17.1 per cent). Vulnerability lines associated with the average risk of falling into poverty of income decile groups 1, 2, 3 and 4 are below the proposed high vulnerability line.

Figure 3 shows the income distribution and the two vulnerability line cut-offs that create and classify three groups within the distribution according to their degree of vulnerability, that is: high, moderate and low. The figure shows the size of each group within the income distribution, providing clear guidance to prioritise social policies tailored to each group. This is, policies aimed to prevent that those facing high vulnerability fall into poverty again, and support those experiencing moderate vulnerability so they can enter the income secure middle-class instead of moving backwards to face either high vulnerability or poverty.

Figure 3: Income distribution by degrees of vulnerability to poverty in Chile



Source: Author's calculations using the P-CASEN 2006-2009.

Notes: The dark dots indicate the association between the probability of falling into poverty in the next year and the income level for the deciles of the income distribution.

Furthermore, Figure 3 shows the average risk of falling into poverty for the decile groups of the income distribution, which is associated with the average household income in each decile group. This information is also relevant for the design of social policies that aim to focus resources and obtain a greater impact within these three groups with different levels of vulnerability to poverty. For example, Figure 3 shows that almost one-third of the population face a high vulnerability, where decile groups 2 and 3 have a probability of falling into poverty of more than a 25 per cent, while decile group 4 has a probability of entering poverty of 20 per cent. Knowing this gradient enables the design of differentiated policies for each of the groups.

Other advantage of the strategy that I have designed is that it shows the relationship between the household per capita income and the probability of falling into poverty for each of the three groups separately. This allows for using the point estimates of the poverty transition equations to examine how the predicted probabilities of poverty entry vary for individuals and households in each group with different combinations of characteristics.

In order to ascertain whether the three groups identified according to their level of vulnerability to poverty differ from each other, I estimate a three-group mean comparison to test if there is a significant difference between the characteristics of: i) the poor and those who are highly vulnerable, ii) those who are highly vulnerable and those who are moderately vulnerable, and iii) households that are moderately vulnerable and those who show low vulnerability.

Table 5 shows that the differences between the four groups are broad and statistically significant for most of the variables, particularly those related to the type and structure of the household. However, when comparing the moderately vulnerable with those less vulnerable to poverty, the variables related to the labour status of the head of household and the number of workers in the household are not statistically significant, whereas the variable that reports on whether the head of household has a university degree presents the greatest average difference between the two groups. This suggests that for a household to transit to a low risk of falling into poverty (i.e., to become income-secure middle class) the number of household members that are working is less relevant than the head of household or the partner having a university degree.

Table 5: Characteristics of the household in the last year (t-1) by degrees of vulnerability to poverty in Chile (*Percentage of household and three-group mean-comparison t-test*)

Variables	Vulnerability to poverty classification				Level of significance of differences between two groups (ref. 95 %)		
	Poverty	High vulnerability	Moderate vulnerability	Low vulnerability	Poverty & High vul.	Hig vul. & Mod. vul.	Mod. vul. & Low. vul.
Household head characteristics							
Female	0.28	0.25	0.28	0.33	0.047	0.012	0.004
Age	44.7	47.0	49.2	49.5	0.000	0.000	0.527
Education: Primary school	0.50	0.42	0.33	0.19	0.000	0.000	0.000
Education: Secondary school	0.43	0.51	0.57	0.53	0.000	0.001	0.063
Education: University degree	0.01	0.02	0.08	0.27	0.199	0.000	0.000
Labour status: Formal employed	0.53	0.67	0.75	0.77	0.000	0.000	0.142
Labour status: Informal employed	0.22	0.15	0.09	0.09	0.000	0.000	0.770
Labour status: Unemployed	0.04	0.02	0.00	0.01	0.002	0.000	0.384
Labour status: Inactive	0.18	0.15	0.13	0.12	0.034	0.085	0.368

HH head's partner characteristics							
Age	41.7	43.6	45.6	46.1	0.000	0.000	0.151
Education: Primary school	0.31	0.27	0.21	0.10	0.04	0.000	0.000
Education: Secondary school	0.24	0.31	0.35	0.35	0.000	0.004	0.801
Education: University degree	0.00	0.01	0.02	0.12	0.018	0.000	0.000
Labour status: Formal employed	0.04	0.11	0.20	0.28	0.000	0.000	0.000
Labour status: Informal employed	0.04	0.07	0.07	0.06	0.002	0.890	0.283
Labour status: Unemployed	0.04	0.03	0.02	0.01	0.100	0.028	0.024
Labour status: Inactive	0.44	0.39	0.31	0.22	0.026	0.000	0.000
Household characteristics							
Household type: Couple without children	0.11	0.21	0.30	0.36	0.000	0.000	0.000
Household type: Single without children	0.05	0.08	0.13	0.18	0.032	0.000	0.000
Household type: Couple with children	0.60	0.52	0.39	0.28	0.000	0.000	0.000
Household type: Single with children	0.22	0.16	0.10	0.05	0.001	0.000	0.000
Household type: Lone person	0.02	0.03	0.07	0.13	0.073	0.000	0.000
Number of persons	4.7	4.3	3.8	3.2	0.000	0.000	0.000
Number of children < 15	1.7	1.2	0.7	0.5	0.000	0.000	0.000
Number of workers	1.0	1.4	1.7	1.7	0.000	0.000	0.871
Housing: Own housing (no mortgage)	0.50	0.59	0.62	0.59	0.000	0.071	0.033
Housing: Own housing, mortgage	0.04	0.09	0.15	0.21	0.000	0.000	0.000
Housing: Rent	0.11	0.12	0.11	0.13	0.485	0.191	0.032
Housing: Subsidized or rent free	0.35	0.20	0.12	0.07	0.000	0.000	0.000
Rural	0.24	0.18	0.10	0.06	0.000	0.000	0.000
Regions: 1st, 2nd, 3rd and 4th	0.13	0.12	0.13	0.13	0.322	0.395	0.781
Regions: 5th, 6th, 7th, 8th, 9th and 10th	0.63	0.57	0.47	0.42	0.006	0.000	0.002
Regions: 11th and 12th	0.00	0.01	0.01	0.02	0.002	0.091	0.371
Regions: 13th	0.24	0.30	0.38	0.43	0.001	0.000	0.006
Nº household	922	1,667	2,267	1,391			

Source: Author's calculations using the P-CASEN 2006-2009.

Notes: Reference year is 2006. Cross-sectional weights are used.

Finally, I present an application of the estimated vulnerability lines to the cross-sectional CASEN data surveys from 2000 to 2017. Table 6 shows the evolution of the degrees of vulnerability to poverty in Chile from the period studied. Using the two vulnerability lines for the poverty line of $5.5 dollars pppd (2011 PPP) allows to study the evolution of changes in the economic well-being of Chilean households groups with different levels of vulnerability. In 2000, poor, highly vulnerable and moderately vulnerable households represented 75 per cent of the population in equal proportions (25 per cent each). However, during the period studied, the evolution of these groups was very different. While poor households drastically decreased, reaching 2.8 per cent in 2017, households with high vulnerability also declined. However, in 2017 they still represented the 12 per cent of the population. In the case of moderate households, the trend was the opposite, reaching 34.9 per cent in 2017. Until 2013, the share of this group of households was the highest in the Chilean population. In 2017 alone, the economically secure middle class with a low risk of falling into poverty appeared as the largest group in Chilean society, with 42.5 per cent.

Table 6: Evolution of the degrees of vulnerability to poverty in Chile from 2000 to 2017

Degrees of vulnerability to poverty	2000	2003	2006	2009	2011	2013	2015	2017
Poverty	25.8	21.0	16.5	12.6	10.8	5.2	3.8	2.8
High vulnerable	24.5	25.4	24.7	23.8	22.4	18.0	15.1	12.0
Moderate vulnerable	25.9	28.2	30.5	32.6	34.1	37.3	36.9	34.9
Low vulnerable (middle class)	19.3	20.9	23.7	25.7	26.4	33.0	37.8	42.5
Not vulnerable	4.5	4.5	4.6	5.4	6.3	6.5	6.4	7.9
Total	100	100	100	100	100	100	100	100

Source: Author's calculations using cross-sectional CASEN data surveys.

Results in Table 6 differ from the World Bank's diagnosis of Chile using a single line of economic vulnerability (World Bank, 2021). Using the 13 dollars pppd (2011 PPP) vulnerability line for the 5 dollars pppd (2011 PPP) poverty line, they estimated that Chile reduced vulnerable households from 36.7 per cent in 2009 to 28.5 per cent in 2017, while the middle class grew from 35.8 per cent to 58.8 per cent during the same period (World Bank, 2021). These differences in analysing the same data open the discussion on the importance of incorporating degrees of vulnerability to poverty in the design of social protection policies and defining what is understood as a middle-class household from an economic security perspective.

6. Implications of using vulnerability lines based on a low-income dynamics approach

Schotte et al. (2018), in their work, question the use of income cut-offs to classify vulnerable-to-poverty households. Using income thresholds based on a poverty dynamics approach to identify degrees of vulnerability to poverty allows to discuss the critical point raised by these authors utilising the concept of 'plausible inconsistency'. It helps to understand better the cases that do not meet the assumption of monotonicity between the expected income and the expected probability of entering poverty. Thus, in this section, I first show the importance of household characteristics on the predicted probabilities of entering poverty for different combinations of household types for stylised examples. Second, I define the concept of 'plausible inconsistency', illustrate some instances, and discuss its implications. Third, I perform a prediction comparison for different vulnerability cut-offs.

Predicted probabilities of poverty entry for different household characteristics

In order to demonstrate the scope of the analysis enabled by this approach I have estimated for twelve family types, the household income and their probability of falling into poverty along with their non-poverty spell duration.[26] To carry out this exercise I used the estimated points of the parameters of the model that controls for the selection biases associated with initial condition and attrition (Table 2).

Table 7 shows the results of the stylised families associated with the household welfare classification obtained from the low and high vulnerability lines. The households are listed according to their position in the income distribution. The reference household type (Case 1) is at the upper end of the income distribution with an estimated income of $81.1 dollars pppd in 2011 PPP, and a risk of falling into poverty close to zero; it is classified in the category "affluent professional".

In column 1 of Table 7, the characteristics of households that change from one case to another are detailed. These changes are related to an increase in the probability of falling into poverty. In this way, the table offers a depiction of the household types that fit into the different classifications. It also provides information on the households whose income is close to the vulnerability lines used to distinguish one group from another.

For example, a household that is classified as 'income-secure middle class' has an income of $25.1 dollars pppd in 2011 PPP and a risk of entering into poverty of 7.5 per cent. This case, which corresponds to Case 5 in Table 4, is a household formed by a couple with one child aged over 15 years. The head of the household is a 45-year-old male, who has completed university education and is formally employed. His partner is 40 years old, has completed secondary school, and is inactive. They rent their house in an urban area in the capital city.

Of particular interest is Case 8 in Table 7. This household differs from Case 5 because it has two children, the head of the household has only secondary education, and his partner works in the informal sector of the economy. The estimated household income is $15.3 dollars pppd in 2011 PPP and the probability of entering into poverty in the next year is 13.7 per cent. Following the current criterion of the World Bank (2018) this household would be considered middle-class

[26] Based on the assumption that the relevant processes occur under a steady state equilibrium, it is possible to estimate the length of time spent as non-poor. I use the median non-poverty duration defined as log(0.5)/log(1-e_{it}) (Boskin & Nold, 1975; Cappellari & Jenkins, 2004).

despite having a risk of falling into poverty of over 10 per cent. Under the criteria I propose, using two lines of vulnerability, this household would be classified as moderately vulnerable.

Table 7: Estimates of predicted probability of falling into poverty and durations for stylised households

Household Characteristics (types)	Household per capita income a day (in 2011 PPP)	Household welfare classification	Upper middle-income countries poverty line ($5.5 pppd in 2011 PPPP)	
			Poverty entry rate (e_{it})	Non-poverty spell duration in years (median)
Case 1: Couple with one child aged over 15 years. The head of the household is a 50-year-old male. His partner is 45 years old. Both have completed university education and, are employed in formal work. They reside in their own housing (paying a mortgage) in an urban area in the Capital city.	81.1	Affluent professionals	0.001	692.8
Case 2: Case 1 except child is under 15 years old. The head of the household is 45 years old and his partner is 40 years old.	64.3	At the edge of income-secure middle class	0.003	230.7
Case 3: Case 2 except head of household's partner has only completed secondary school.	46.4	Income-secure middle class	0.010	69.0
Case 4: Case 3 except they rent their house.	35.4	Income-secure middle class	0.042	16.2
Case 5: Case 4 except head of household's partner is inactive.	25.1	Income-secure middle class	0.075	8.9
Case 6: Case 5 except head of household has only completed secondary school and his partner is employed in formal work.	19.9	At the edge of moderate vulnerability	0.080	8.3
Case 7: Case 6 except they have one additional child aged over 15 years old in the household.	16.9	Moderate vulnerability	0.088	7.5
Case 8: Case 7 except head of household's partner is employed in informal work.	15.3	Moderate vulnerability	0.137	4.7
Case 9: Case 8 except head of household's partner is inactive.	12.0	At the edge of high vulnerability	0.148	4.3
Case 10: Case 9 except household is employed in informal work.	9.9	High vulnerability	0.264	2.3
Case 11: Case 10 except head of household's partner has only completed primary school.	7.9	High vulnerability	0.312	1.9
Case 12: Case 11 except the head of household is female and her partner is unemployed.	6.1	At the edge of poverty	0.348	1.6

Source: Author's calculations using the P-CASEN 2006-2009.

Note: Estimates are based on expressions 19 and 23, point estimates from Table 2.

Plausible inconsistencies between predicted income household and predicted probability poverty entry

As I have mentioned, my approach has an implicit assumption of monotonicity between the base period household income (among non-poor) and the probability of poverty entry, that is, higher income implies a lower probability of falling into poverty. However, when applying vulnerability lines to distinguish the degree of vulnerability to poverty, I risk making misclassification errors because there are cases where that assumption is not met. I argue that these cases can be seen as 'plausible inconsistencies'. Looking at the variables of the models that predict both the probability to falling into poverty and household income, 'plausible inconsistencies' are found to explain, for example, cases of households that share the same income but face different poverty entry probability, and inversely, households that share the same poverty entry risk but have different incomes.

Table 8: Comparison between different households with the same predicted daily income and with the same probability of falling into poverty in the next year

Household Characteristics	Household per capita income a day (in 2011 PPP)	Upper middle-income countries poverty line $5.5 pppd in 2011 PPPP)	
		Poverty entry probability (e_{it})	Non-poverty spell duration in years (median)
Base Case: Household compound by a couple. They rent their house and reside in an urban area in the Chilean capital city of Santiago.			
1. Two types of household with similar predicted daily income			
Family A: Couple with one child. The head of the household is a 40-year-old male. His partner is 35 years old. Both have completed secondary education. The head of the household is employed in the informal sector of the economy. His partner is also employed in informal work.	15.0	0.232	2.6
Family B: Couple with two children. The head of the household is a 45-year-old male. His partner is 40 years old. Both have completed secondary education. The head of the household is employed in the formal sector of the economy. His partner is employed in informal work.	15.3	0.137	4.7
2. Two types households with a similar probability of falling into poverty in the next year			
Family C: Couple with two children. The head of the household is a 40-year-old male. His partner is 35 years old. Both have completed secondary education and are employed in the formal sector of the economy.	16.2	0.101	6.5
Family D: Couple without children. The head of the household is a 60-year-old male. His partner is 50 years old. Both have completed secondary education. The head of the household is unemployed. His partner is employed in formal work.	11.2	0.101	6.5

Source: Author's calculations using the P-CASEN 2006-2009.

Note: Estimates are based on equations 7 and 11, point estimates from Table 2.

Table 8 illustrates some examples of 'plausible inconsistencies'. I take two households that would be classified as middle-class using the World Bank's vulnerability line, with an income per capita close to $15.0 dollars pppd in 2011 PPP, and compare them (Case A and Case B are shown in the first panel of Table 5). Household A differs in terms of two characteristics from household B: instead of two children they have only one, and the head of the household works in an informal job.

The probability of falling into poverty for Case B (a household of four people) is 13.7 per cent, whereas Case A, though a smaller household, has two members working in the informal sector and shows a probability of falling into poverty that is around double that of Case B. This result should not be surprising because it reflects the economic insecurity of a household with two informal workers, despite the fact that it has fewer members than the other.

Likewise, there are households with the same risk of entering poverty and different income levels. In panel B of Table 8, an example of this is shown. Case C and Case D describe households with different characteristics, namely, number of children, age of the couple, size of the household and number of people working. However, despite their level of per capita income being different, they have the same risk of entering poverty.

The existence of 'plausible inconsistencies' in the classification of the households according to degrees of vulnerability to poverty connects with the discussion posed by Schotte et al. (2018), who strongly question the use of income cut-offs, proposing instead the use of the poverty entry probability thresholds to classify groups within the income distribution. My results show that non-compliance with the monotonicity assumption between income and risk of falling into poverty may not necessarily be seen as a classification problem. Both outcomes are plausible to be used to classify groups with different risks of falling into poverty. However, it could be argued that the use of vulnerability lines could have a greater problem of 'accuracy' that using poverty risk thresholds (Celidoni, 2013; Hohberg et al., 2018).

Comparison of prediction for different vulnerability cut-offs

Table 9 shows the results of the comparison of prediction for both vulnerability cut-offs.[27] It also compares the use of degrees of vulnerability to poverty versus a simple dichotomy of vulnerable versus non-vulnerable. I chose the years 2007 and 2008 to show the households 'situation just

[27] It is important to note that Table 9 is only a prediction comparison, not a comparison of predictive performance. To assess predictive performance, I need a measure of predictive performance (for example, the area under the ROC curve) relative to a reference classification. However, such a reference classification does not exist in this case since the model generates the classification. New research related to the literature on targeting performance should delve into the classification errors caused by different types of cut-offs and the use of more than one vulnerability line.

before the economic crisis and one year after.[28]

Table 9: Comparison of prediction of degrees of vulnerability to poverty and vulnerability to poverty for different vulnerability cut-offs

a) Degrees of vulnerability to poverty

Two vulnerability lines			Two probability cut-offs		
	2008			2008	
2007	Poor	Non-poor	2007	Poor	Non-poor
Highly vulnerable	24.8	75.2	Highly vulnerable	21.3	78.74
Moderately vulnerable	7.8	92.2	Moderately vulnerable	7.7	92.3
Lowly vulnerable (Middle-class)	3.7	96.3	Lowly vulnerable (Middle-class)	1.7	98.4

b) Vulnerability to poverty

One vulnerability line			One probability cut-off		
	2008			2008	
2007	Poor	Non-poor	2007	Poor	Non-poor
Vulnerable	20.0	80.0	Vulnerable	17.9	82.13
Non-vulnerable (Middle-class)	4.3	95.7	Non-vulnerable (Middle-class)	4.4	95.6

Source: Author's calculations using the P-CASEN 2006-2009.

Notes: Estimates are based on balance data between 2007-2008 using survey longitudinal weights. Vulnerability lines derived from the World Bank poverty line ($5.5 dollars pppd in 2011 PPP).

Panel A of Table 9 shows those who were classified with degrees of vulnerability using two vulnerability cut-offs. Using vulnerability lines in the classification, one of four highly vulnerable households fell into poverty. Using probability cut-offs this percentage is lower: 21.3 per cent of highly vulnerable households were poor in 2008. For those who are moderately vulnerable to both types of cut-offs, the percentage was close to 8 per cent. The percentage of households with a low vulnerability that fall into poverty was 3.7 per cent for the income threshold and less than 2 per cent for the poverty risk cut-offs. Using the official poverty line for Chile, the proportion of households that enter into poverty is similar, except among the moderately vulnerable where the percentage that falls into poverty in 2008 is around 10 per cent. See Table 2.

As in Panel A, Panel B of Table 9 shows that the vulnerability line performs better to predict who fell into poverty than the probability cut-off. From the comparison of Panel A with Panel B, it is possible to identify two advantages of classifying households according to their degree of vulnerability rather than a simple dichotomy of vulnerable versus non-vulnerable. First, using two vulnerability lines allows for a better prediction of those who fall into poverty. For instance, it

[28] Although the collapse of the housing bubble in the United States began in 2006, the so-called subprime mortgage crisis began to spread to international markets from October 2007 onward, with 2008 being its worst year (IMF, 2009).

enables to compare the highly vulnerable with those identified as vulnerable using one vulnerability line. Second, it identifies the moderately vulnerable group whose proportion of households that fall into poverty is significantly higher than the percentage among the economically secure (middle class) group.

In the short term, using more than one vulnerability line to identify different non-poor vulnerable groups provides better information to policymakers to design and implement social protection programs to face situations such as an economic crisis. In the long-term, it improves the anti-poverty targeting performance in countries with a weak welfare state and a distribution of income that is markedly displaced to the left around the poverty line.

7. Conclusion

In this paper, I have proposed an empirical framework to identify different degrees of vulnerability to poverty within the income distribution using a poverty dynamics approach. Applying this approach to household data from Chile, I estimate low and high vulnerability lines. This allows the identification of three types of households: those with high, moderate and low vulnerability to poverty. The latter is the income secure middle class. Distinguishing between different types of vulnerability is crucial not only for the design of social policies targeted at families at high risk of poverty, but also for understanding the characteristics of those who experience greater economic stability or security.

Assuming that the economic conditions that determine vulnerability remain unchanged in the future, the thresholds in real income terms can be used to measure the size and evolution of vulnerable groups using cross-sectional household surveys. In the case of Chile,using a poverty line of $5.5 dollars pppd (2011 PPP), high vulnerability households are those with a per capita income between $5.5 and $9.9 (above the poverty line and below the high vulnerability line), moderate vulnerability households are those with a per capita income between $9.9 and $20.0 dollars (between the high and low vulnerability lines), and low vulnerability households - the income secure middle class - are those with a per capita income between $20 and $70 dollars pppd.

My approach proposes a more demanding definition of the middle class than that proposed by the World Bank (between $13.0 and $70.0 pppd (2011 PPP)). This is because it distinguishes two vulnerable groups rather than one: those at high and moderate risk of experiencing poverty in the near future. It is worth noting that the World Bank's vulnerability line and the one I propose have different interpretations. The World Bank's vulnerability line is associated with the risk of all non-poor households falling into poverty, estimated using panel data with long intervals (3 to 5 years). Instead, I propose low and high vulnerability lines associated with the probability of falling into poverty from one year to the next for different groups within the income distribution. The use of a 'one year to the next' criterion not only allows for a more precise identification of vulnerable groups, but can also better serve the implementation of risk management and anti-poverty policies.

The implications of these results are significant. A large proportion of the population that would be classified as middle class according to the World Bank's vulnerability line are households that, according to my approach, face considerable economic insecurity. I would classify them as moderately vulnerable. Based on these findings, I argue that previous research has underestimated how many people in Chile are at risk of falling into poverty and overestimated the growth of the middle class. These sobering conclusions should be of great interest to Chilean policymakers and others in other middle-income countries that use the World Bank's vulnerability line, especially in Latin America.

Vulnerability to poverty lines offer governments a concrete way to improve the targeting of programmes aimed at reducing absolute poverty. The extension of social protection coverage to these new social groups should be accompanied by a comprehensive design of social protection programmes that includes vulnerability to poverty as part of economic welfare measures to assess social progress. In this way, the approach to vulnerability to poverty that I have proposed should play a dual role in targeting and monitoring these new social groups.

Finally, my study lies at the intersection of the interests of several disciplines, in particular economics and sociology. It contributes to the economic literature not only by bridging the gap between the vulnerability to poverty and poverty dynamics approaches, but also by empirically determining the income cut-offs to identify degrees of vulnerability to poverty that go beyond the distinction between vulnerable and non-vulnerable. It also contributes to the discussion on social stratification in sociology, since the approach I propose, based on degrees of vulnerability to poverty, is better adapted to the reality of middle-income countries and to the definitions proposed by this discipline to conceptualise and measure the middle class.

9. References

Alesina, A., & Perotti, R. (1996). Income distribution, political instability, and investment. *European Economic Review*, *40*, 1203–1228.

Atkinson, A., & Brandolini, A. (2013). On the identification of the middle class. In *Income inequality: Economic disparities and the middle class in affluent countries* (pp. 77–100).

Ayllón, S. (2013). Understanding poverty persistence in Spain. *SERIEs - Journal of the Spanish Economic Association*, *4*, 201–233.

Banerjee, A. V., & Duflo, E. (2008). What is Middle Class about the Middle Classes around the World? *Journal of Economic Perspectives*, *22*, 3–28.

Barro, R. J. (2000). Inequality and Growth in a Panel of Countries. *Journal of Economic Growth*, *5*, 5–32.

Bendezú, L., Denis, A., & Zubizarreta, J. R. (2007). *Análisis de la Atrición de la Muestra en la Encuesta Panel CASEN* (p. 22). Santiago: Mimeo, Observatorio Social, Universidad Alberto Hurtado.

Birdsall, N. (2010). *The (Indispensable) Middle Class in Developing Countries; or, The Rich and the Rest, Not the Poor and the Rest* (CGD Working Paper No. 207). Washington DC: Center for Global Development.

Birdsall, N. (2015). Does the Rise of the Middle Class Lock in Good Government in the Developing World? *The European Journal of Development Research*, *27*, 217–229.

Birdsall, N., Graham, C., & Pettinato, S. (2000). *Stuck In The Tunnel: Is Globalization Muddling The Middle Class?* (CGD Working Paper No. 14; p. 37). Washington DC: Center for Global Development.

Birdsall, N., Lustig, N., & Meyer, C. J. (2014). The Strugglers: The New Poor in Latin America? *World Development*, *60*, 132–146.

Boskin, M. J., & Nold, F. C. (1975). A Markov Model of Turnover in Aid to Families with Dependent Children. *The Journal of Human Resources*, *10*, 467–481.

Buddelmeyer, H., & Verick, S. (2008). Understanding the Drivers of Poverty Dynamics in Australian Households*. *Economic Record*, *84*, 310–321.

Cappellari, L., & Jenkins, S. P. (2004). Modelling low income transitions. *Journal of Applied Econometrics*, *19*, 593–610.

Celidoni, M. (2013). Vulnerability to poverty: An empirical comparison of alternative measures. *Applied Economics*, *45*, 1493–1506.

Celidoni, M. (2015). Decomposing Vulnerability to Poverty. *Review of Income and Wealth*, *61*, 59–74.

Ceriani, L. (2018). Vulnerability to poverty: Empirical findings. In C. D'Ambrosio (Ed.), *Handbook of Research on Economic and Social Well-Being*. Edward Elgar Publishing.

Chaudhuri, S. (2003). *Assessing vulnerability to poverty: Concepts, empirical methods and illustrative examples*. New York: Department of Economics, Columbia University.

Chaudhuri, S., Jalan, J., & Suryahadi, A. (2002). *Assessing household vulnerability to poverty from cross-sectional data: A methodology and estimates from Indonesia* (No. 0102–52). New York: Department of Economics, Columbia University.

Chauvel, L. (2018). *A new step in the understanding of extreme inequality dynamics: Chile comes with 12 waves 1990- 2015*. LIS Newsletter, Issue No. 6.

Chiwaula, LevisonS., Witt, R., & Waibel, H. (2011). An Asset-Based Approach to Vulnerability: The Case of Small-Scale Fishing Areas in Cameroon and Nigeria. *The Journal of Development Studies*, *47*, 338–353.

Christiaensen, L. J., & Subbarao, K. (2005). Towards an Understanding of Household Vulnerability in Rural Kenya. *Journal of African Economies*, *14*, 520–558.

Cingano, F. (2014). *Trends in Income Inequality and its Impact on Economic Growth* (OECD Social, Employment and Migration Working Papers No. 163). OECD Publishing.

Dang, H.-A., & Dabalen, A. L. (2018). Is Poverty in Africa Mostly Chronic or Transient? Evidence from Synthetic Panel Data. *The Journal of Development Studies*, 1–21.

Dang, H.-A., & Lanjouw, P. (2017). Welfare Dynamics Measurement: Two Definitions of a Vulnerability Line and Their Empirical Application. *Review of Income and Wealth*, *63*, 633–660.

Davis, J. C., & Huston, J. H. (1992). The shrinking middle-income class: A multivariate analysis. *Eastern Economic Journal*, *18*, 277–285.

Denis, A., Prieto, J., & Zubizarreta, J. R. (2007). Dinámica de la Pobreza en Chile: Evidencias en los Años 1996, 2001 y 2006. *Persona y Sociedad*, *21*, 9–30.

Dercon, S. (Ed.). (2005). *Insurance Against Poverty*. Oxford University Press.

Duan, N. (1983). Smearing Estimate: A Nonparametric Retransformation Method. *Journal of the American Statistical Association*, *78*, 605–610.

Easterly, W. (2001). The Middle Class Consensus and Economic Development. *Journal of Economic Growth*, *6*, 317–335.

Estache, A., & Leipziger, D. (2009). *Stuck in the Middle: Is Fiscal Policy Failing the Middle Class?* Brookings Institution Press.

Feeny, S., & McDonald, L. (2016). Vulnerability to Multidimensional Poverty: Findings from Households in Melanesia. *The Journal of Development Studies*, *52*, 447–464.

Ferreira, F., Messina, J., Rigolini, J., López-Calva, L., Lugo, M. A., & Vakis, R. (2013). *Economic Mobility and the Rise of the Latin American Middle Class*. Washington, DC: World Bank.

Gallardo, M. (2013). Using the downside mean-semideviation for measuring vulnerability to poverty. *Economics Letters*, *120*, 416–418.

Gallardo, M. (2018). Identifying vulnerability to poverty: A critical survey. *Journal of Economic Surveys*, *32*, 1074–1105.

Gallardo, M. (2019). Measuring Vulnerability to Multidimensional Poverty. *Social Indicators Research*.

Goldthorpe, J. H., & McKnight, A. (2006). The economic basis of social class. In S. L. Morgan, D. B. Grusky, & G. S. Fields (Eds.), *Mobility and inequality: Frontiers of research from sociology and economics* (pp. 109–136). Palto Alto, CA: Stanford University Press.

Gornick, J. C., & Jäntti, M. (2014). *Income Inequality: Economic Disparities and the Middle Class in Affluent Countries*. Stanford University Press.

Günther, I., & Harttgen, K. (2009). Estimating Households Vulnerability to Idiosyncratic and Covariate Shocks: A Novel Method Applied in Madagascar. *World Development*, *37*, 1222–1234.

Günther, I., & Maier, J. K. (2014). Poverty, Vulnerability, and Reference-Dependent Utility. *Review of Income and Wealth*, *60*, 155–181.

Hoddinott, J., & Quisumbing, A. (2010). Methods for microeconometric risk and vulnerability assessment. In R. Fuentes-Nieva & P. A. Seck (Eds.), *Risk, Shocks, and Human Development* (pp. 62–100). London: Palgrave Macmillan.

Hohberg, M., Landau, K., Kneib, T., Klasen, S., & Zucchini, W. (2018). Vulnerability to poverty revisited: Flexible modeling and better predictive performance. *The Journal of Economic Inequality*, *16*, 439–454.

IMF (Ed.). (2009). *Crisis and recovery*. Washington, DC: International Monetary Fund.

Jenkins, S. P. (2011). *Changing Fortunes: Income Mobility and Poverty Dynamics in Britain*. New York: Oxford University Press.

Jolliffe, D., & Prydz, E. (2016). Estimating international poverty lines from comparable national thresholds. *Journal of Economic Inequality*, *14*, 185–198.

Kanabar, R. (2017). In or out? Poverty dynamics among older individuals in the UK. *Journal of Pension Economics and Finance*, *16*, 509–553.

Kharas, H. (2017). *The unprecedented expansion of the global middle class* (No. 100; p. 32). Washington, DC: Brookings Institution.

Klasen, S., & Waibel, H. (2015). Vulnerability to Poverty in South-East Asia: Drivers, Measurement, Responses, and Policy Issues. *World Development*, *71*, 1–3.

Larrañaga, O. (2009). *Inequality, poverty and social policy: Recent trends in Chile* (No. 85). OECD Social, Employment and Migration,.

Larrañaga, O., & Rodríguez, M. E. (2015). Desigualdad de ingresos y pobreza en Chile: 1990 a 2013. In D. Contreras & O. Larrañaga (Eds.), *Las nuevas políticas de protección social en Chile* (pp. 251–285). Santiago de Chile: Uqbar Editores.

Ligon, E., & Schechter, L. (2003). Measuring Vulnerability. *The Economic Journal*, *113*, C95–C102.

López-Calva, L., & Ortiz-Juarez, E. (2014). A vulnerability approach to the definition of the middle class. *Journal of Economic Inequality*, *12*, 23–47.

Lipton, M. (1983). Poverty, undernutrition, and hunger (World Bank Staff Working Paper 597). World Bank.

Lynn, P., Zubizarreta, J. R., & Castillo, E. (2007). *Panel CASEN Survey: Sample Design*. Santiago, Chile: Observatorio Social, Universidad Alberto Hurtado.

Maldonado, L., Prieto, J., & Lay, S. L. (2016). *Las dinámicas de la pobreza en Chile durante el periodo 2006-2009* (No. 87). Santiago, Chile: Centro UC de Políticas Públicas.

Maldonado, L., & Prieto, J. (2015). Determinantes de la dinámica de la pobreza en Chile y el rol de la persistencia temporal: Análisis de la Encuesta Panel Casen 2006-2009 con métodos de historia de eventos. *Economía y Política*, 2(2), 5–39.

MDS. (2018). *Pobreza y Distribución de Resultados: Resultados CASEN 2017*. Santiago: Ministerio de Desarrollo Social, Gobierno de Chile.

Mideplan. (2010). *CASEN 2009. Encuesta de Caracterización Socioeconómica Nacional*. Ministerio de Desarrollo Social, Gobierno de Chile.

Milanovic, B., & Yitzhaki, S. (2002). Decomposing World Income Distribution: Does The World Have A Middle Class? *Review of Income and Wealth*, *48*, 155–178.

Morduch, J. (1994). Poverty and vulnerability. *American Economic Review*, *84*, 221.

Neilson, C., Contreras, D., Cooper, R., & Hermann, J. (2008). The Dynamics of Poverty in Chile. *Journal of Latin American Studies*, *40*, 251–273.

Observatorio Social. (2011a). *Documento Metodológico Encuesta Panel Casen 2009* (No. 4; p. 126). Santiago: Universidad Alberto Hurtado.

Observatorio Social. (2011b). *Encuesta Panel Casen 2006, 2007, 2008, 2009. Imputación de Ingresos. Informe V1*. Universidad Alberto Hurtado.

Observatorio Social. (2011c). *Encuesta Panel Casen 2009. Informe N° 4. Documento Metodológico Encuesta Panel Casen 2009*. Universidad Alberto Hurtado.

OECD. (2018a). *A Broken Social Elevator? How to Promote Social Mobility*. OECD Publishing.

OECD. (2018b). *OECD Economic Surveys: Chile 2018*. Paris: OECD Publishing.

Paredes, R., Prieto, J., & Zubizarreta, J. R. (2006). *Attrition in Longitudinal Data and Income Mobility in Chile*. Santiago de Chile: Mimeo, Observatorio Social, Universidad Alberto Hurtado.

PNUD. (2017). *Desiguales: Orígenes, cambios y desafíos de la brecha social en Chile*. Santiago de Chile: Programa de las Naciones Unidas para el Desarrollo.

Pons, V., Mullins, W., Masko, J., Lobb, A., & Tella, R. D. (2020). *Unrest in Chile* (No. 720–033). Harvard Business School Case.

Pritchett, L., Suryahadi, A., & Sumarto, S. (2000). *Quantifying vulnerability to poverty: A proposed measure, applied to Indonesia* (No. 2437). Washington, DC: World Bank.

Ravallion, M. (1996). Issues in Measuring and Modelling Poverty. *The Economic Journal*, *106*, 1328–1343.

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

Santos Silva, J. M. C., & Tenreyro, S. (2006). The Log of Gravity. *The Review of Economics and Statistics*, *88*, 641–658.

Schotte, S., Zizzamia, R., & Leibbrandt, M. (2018). A poverty dynamics approach to social stratification: The South African case. *World Development*, *110*, 88–103.

Skoufias, E., & Quisumbing, A. R. (2005). Consumption Insurance and Vulnerability to Poverty: A Synthesis of the Evidence from Bangladesh, Ethiopia, Mali, Mexico and Russia. *European Journal of Development Research*, *17*, 24–58.

Stampini, M., Robles, M., Sáenz, M., Ibarrarán, P., & Medellín, N. (2016). Poverty, vulnerability, and the middle class in Latin America. *Latin American Economic Review*, *25*.

Suryahadi, A., & Sumarto, S. (2003). Poverty and Vulnerability in Indonesia Before and After the Economic Crisis. *Asian Economic Journal*, *17*, 45–64.

Tezanos, S., & Sumner, A. (2016). Is the 'Developing World' Changing? A Dynamic and Multidimensional Taxonomy of Developing Countries. *The European Journal of Development Research*, *28*, 847–874.

Torche, F. (2005). Unequal but fluid: Social mobility in Chile in comparative perspective. *American Sociological Review*, *70*, 422–450.

Wietzke, F.-B., & Sumner, A. (2018). The Developing World's "New Middle Classes": Implications for Political Research. *Perspectives on Politics*, *16*, 127–140.

World Bank. (2012). *Well Begun, Not yet Done-Vietnam's Remarkable Progress on Poverty Reduction and the Emerging Challenges*. Hanoi: World Bank.

World Bank. (2018). LAC Equity Lab: Poverty. http://www.worldbank.org/en/topic/poverty/lac-equity-lab1/poverty

World Bank. (2021). ¿Cómo mejorar la protección a la clase media chilena? https://www.bancomundial.org/es/events/2021/03/31/webinar-c-mo-mejorar-la-protecci-n-a-la-clase-media-en-chile

Zizzamia, R., Schotte, S., Leibbrandt, M., & Vimal Ranchhod. (2016). *Vulnerability and the middle class in South Africa* (SALDRU Working Paper No. 188). NIDS.

10. Appendices

A. Tables

Table A.1: Model estimates of poverty entry rates, initial poverty status and survey retention, Chile (2006-2009)

Variables (measured at *t*)	Poverty entry: Poor at t+1 | Non-poor at t			Poverty status at t			Survey retention		
	Coefficient		Std. Dev.	Coefficient		Std. Dev.	Coefficient		Std. Dev.
Household head characteristics									
Female	0.028		(0.021)	0.098	***	(0.021)	-0.046	**	(0.021)
Age	-0.006	***	(0.001)	-0.015	***	(0.001)	0.001		(0.001)
Education: Ref. Secondary school									
Primary school	0.129	***	(0.019)	0.3	***	(0.019)	0.133	***	(0.020)
University degree	-0.356	***	(0.041)	-0.717	***	(0.058)	-0.163	***	(0.029)
Labour status: Ref. Formal employed									
Informal employed	0.394	***	(0.025)	0.6	***	(0.024)	0.011		(0.025)
Unemployed	0.171	*	(0.102)	1.086	***	(0.055)	-0.004		(0.076)
Inactive	0.007		(0.042)	0.22	***	(0.036)	-0.006		(0.037)
HH head's partner characteristics									
Age	-0.009	***	(0.001)	-0.002	*	(0.001)	-0.001		(0.001)
Education: Ref. Secondary school									
Primary school	0.228	***	(0.023)	0.281	***	(0.025)	0.045	*	(0.024)
University degree	-0.433	***	(0.074)	-0.872	***	(0.176)	-0.056		(0.042)
Labour status: Ref. Formal employed									
Informal employed	0.283	***	(0.035)	0.306	***	(0.045)	-0.004		(0.035)
Unemployed	0.08		(0.054)	0.295	***	(0.050)	0.227	***	(0.057)
Inactive	0.124	***	(0.023)	0.094	***	(0.025)	0.046	**	(0.022)
Household characteristics									
Household type: Ref. Couple without children									
Single without children	0.154	***	(0.036)	0.145	***	(0.041)	-0.033		(0.033)
Couple with children	0.073	**	(0.029)	0.189	***	(0.033)	0.030		(0.027)
Single with children	0.303	***	(0.035)	0.493	***	(0.038)	0.155	***	(0.035)
Lone person	0.031		(0.067)	-0.052		(0.075)	-0.198	***	(0.056)
Number of persons	0.038	***	(0.008)	0.251	***	(0.008)	-0.028	***	(0.007)
Number of children < 15	0.133	***	(0.013)	0.114	***	(0.013)	0.065	***	(0.012)
Number of workers	-0.162	***	(0.014)	-0.959	***	(0.019)	0.009		(0.010)
Housing: Ref. Own housing (mortgage)									
Own housing, mortgage	-0.307	***	(0.032)	-0.315	***	(0.038)	-0.149	***	(0.025)
Rent	0.233	***	(0.026)	0.43	***	(0.028)	-0.364	***	(0.024)
Subsidized or rent free	0.212	***	(0.024)	0.716	***	(0.021)	-0.075	***	(0.023)
Rural	0.079	***	(0.023)	0.127	***	(0.024)	0.094	***	(0.026)
Regions: Ref. 13th									
1st, 2nd, 3rd and 4th	0.054	**	(0.026)	0.017		(0.029)	0.034		(0.025)
5th, 6th, 7th, 8th, 9th and 10th	0.13	***	(0.019)	0.214	***	(0.021)	0.147	***	(0.018)
11th and 12th	-0.133	***	(0.051)	-0.299	***	(0.059)	0.271	***	(0.055)
Time (*t*): Ref. 2007									
2008	0.145	***	(0.019)	-0.178	***	(0.018)			
2009	-0.101	***	(0.020)	0.012		(0.018)			
Individual characteristics (Exclusion restrictions)									
Mother education: Ref. No schooling									
Primary school				-0.064	*	(0.033)			

Secondary school	-0.142	***	(0.047)			
University degree	-0.227	**	(0.103)			
Type of work done by mother: Ref. Self-employed						
Employership	-0.054		(0.123)			
Paid employment	0.094	**	(0.033)			
Non-employment	0.011		(0.028)			
Father education: Ref. No schooling						
Primary school	0.009		(0.034)			
Secondary school	-0.099	**	(0.045)			
University degree	-0.015		(0.098)			
Type of work done by father: Ref. Self-employed						
Employership	-0.036		(0.073)			
Paid employment	-0.063	***	(0.026)			
Non-employment	-0.022		(0.108)			
Original sample member				0.509	***	(0.056)
Constant	-1.345	***	(0.063)	0.602	***	(0.081)
Log-pseudolikelihood	-52,691.112					
Wald chi-square (d.f. = 131)	487,629.326 (p<0.000)					
Number of observations (person-waves)	65,205					

Source: Author's calculations using the P-CASEN 2006-2009.

Notes: Model used the Chilean official poverty line ($6.41 dollars pppd in 2011 PPP). Robust standard errors clustered at the individual level. Simulated pseudo maximum likelihood estimation with 250 random draws. *** significance at 1 percent; ** significance at 5 percent; * significance at 1 percent.

Table A.2: Comparison of predictive performance between degree vulnerability to poverty and vulnerability to poverty for different vulnerability cut-offs

a) Degree vulnerability to poverty					
Two vulnerability lines			Two probability cut-offs		
	2008			2008	
2007	Poor	Non-poor	2007	Poor	Non-poor
Highly vulnerable	25.4	74.6	Highly vulnerable	21.2	78.8
Moderately vulnerable	10.7	89.3	Moderately vulnerable	10.1	89.1
Lowly vulnerable (Middle-class)	3.3	96.7	Lowly vulnerable (Middle-class)	0.9	99.1
b) Vulnerability to poverty					
One vulnerability line			One probability cut-off		
	2008			2008	
2007	Poor	Non-poor	2007	Poor	Non-poor
Vulnerable	19.9	80.1	Vulnerable	17.6	82.4
Non-vulnerable (Middle-class)	4.6	95.4	Non-vulnerable (Middle-class)	5.2	94.8

Source: Author's calculations using the P-CASEN 2006-2009.

Notes: Estimates are based on balance data between 2007-2008 using survey longitudinal weights. Vulnerability lines derived from the Chilean official poverty line ($6.41 dollars pppd in 2011 PPP).