Assessing the Predictive Power of Vulnerability Measures: Evidence from Panel Data for Argentina and Chile[1]

Marcelo Bérgolo
CEDLAS, Universidad Nacional de La Plata,
and IECON, Universidad de la República*
Guillermo Cruces
CEDLAS, Universidad Nacional de La Plata, and CONICET†[2]
Andrés Ham
University of Illinois at Urbana Champaign‡

This article carries out a validation exercise of vulnerability measures as predictors of poverty at the aggregate and micro levels based on short and long term panel data for Argentina and Chile. It then compares their performance to that of deprivation indicators. The main findings indicate that while vulnerability measures are good predictors of poverty in the aggregate, the same does not occur at household level. These results imply that while useful, vulnerability estimates require incorporating shocks to attenuate biased estimates if they are to be used for targeting purposes.

Keywords: vulnerability, poverty, targeting, Argentina, Chile
JEL Classifications: D31, I32, I38

Introduction

Most policy interventions in the developing world are guided by the fundamental objective of reducing poverty. Policies are designed to tackle the different characteristics, causes and manifestations of this multi-faceted and multi-dimensional problem. In fact, some of the largest policy initiatives in Latin America consist of income safety nets and emergency and conditional cash transfer programs, which are aimed at reducing present deprivation and preventing its persistence.

While some dimensions of poverty are clearly inter-temporal, most distributive analysis and policy design processes in Latin America are based on cross-sectional

**Address for Correspondence*: Instituto de Economía, Facultad de Ciencias Económicas y Adminimstraci'on, Universidad de la República, Joaquín Requena 1375, Montevideo, Uruguay. Email: mbergolo@iecon.ccee.edu.uy.

†*Address for Correspondence*: Centro de Estudios Distributivos, Laborales y Sociales, Facultad de Ciencias EconÃşmicas, Universidad Nacional de La Plata. Calle 6 entre 47 y 48, 5to. piso, oficina 516, (1900) La Plata, Argentina. Email: gcruces@cedlas.org

‡*Address for Correspondence*: 214 David Kinley Hall, 1407 W. Gregory, Urbana, Illinois, 61801, MC-707. Email: hamgonz2@illinois.edu

data and estimates. For instance, poverty profiles routinely identify households considered to be in a precarious or vulnerable state. This reliance on ex-post outcomes has been subject to an in-depth critique in the literature dealing with vulnerability (see, for instance, Chaudhuri, 2003; Dercon, 2006; Hoddinott and Quisumbing, 2008).

These and other studies suggest that the risk of being poor and the actual state of poverty are two related, but separate phenomena. At any given time, a number of non-poor households may be at high risk of falling into poverty in a future period. In this instance, these households would be considered vulnerable as opposed to non-poor. At the same time, there may also be households below the poverty line, which are not vulnerable in this sense and where their observed poverty status reflects a temporary deprivation. Given these varied scenarios, it is necessary to distinguish between vulnerability and the current state of deprivation. This classification gains greater relevancy in light of recent policy developments; for example, a number of Latin American countries have established poverty alleviation strategies and conditional cash transfer programs, the designs of which would greatly benefit from effectively predicting future poverty (for instance, in terms of targeting).

This article addresses the different natures of vulnerability and poverty and empirically estimates the first using cross-sectional data. The discussion does not focus on these calculations, however, but on their effectiveness, defined as the predictive power to forecast future poverty states. In particular, the analysis uses panel data from Argentina and Chile to carry out this task by comparing estimated vulnerability to actual realised poverty states in future periods. In addition, the study also tests how vulnerability estimates perform in this sense against a series of deprivation indicators. The exercise is grounded in a clear policy motivation: good predictive power would make vulnerability measures a superb targeting tool.

The findings presented encompass a series of contributions with respect to the performance of vulnerability measures as predictors of future poverty. Firstly, the analysis assesses how well vulnerability predicts poverty at the aggregate (or national) level. Secondly, the discussion focuses on how effectively the estimates predict whether a specific household will be poor in the future, and quantifies misclassifications (specifically, poor households classified as not vulnerable in the previous period, and non-poor households originally classified as vulnerable). While previous validation exercises concentrated on aggregate vulnerability and poverty levels (Zhang and Wan, 2009), the discussion here argues that the usefulness of vulnerability measures for social policy depends on how well they can identify household-specific rather than aggregate outcomes, in particular among those below and close to the poverty line. Finally, while Chaudhuri *et al.* (2002), Chaudhuri (2003) and Zhang and Wan (2009) based their exercises on short run panels, this study provides evidence from both short run (Argentina) and long run (Chile) panels.

The remainder of this study is organised as follows: the next section presents a conceptual and methodological discussion of vulnerability measures and reviews recent developments in the literature; a following section describes the data sources used to conduct the analysis and establishes the empirical strategy; two subsequent sections present the results of the validation exercise, and a next section compares the predictive power of vulnerability measures with respect to a selection of deprivation indicators; a final section offers some conclusions.

Measuring Vulnerability

Approaches to vulnerability measurement

In abstract terms, vulnerability can be defined as the threat to welfare at a future date. This threat stems from one of two factors: high levels of welfare variability or systematically low levels of welfare.

There are three main approaches to identify the vulnerable, which hinge on how welfare is measured. The first approach assesses vulnerability as expected poverty (VEP). This line of research seeks to estimate the probability that welfare may fall below a minimum expected standard of living in the future (Chaudhuri *et al.*, 2002). The second approach quantifies vulnerability as low expected utility (VEU). This alternative arose because of concerns with the VEP method, which is assumed to omit important issues which VEU incorporates (see Ligon and Schechter, 2003).[3] Finally, the last approach is vulnerability as uninsured exposure to risk (VER). This method, unlike the previous two, concentrates on observing past outcomes rather than predicting future welfare (Tesliuc and Lindert, 2004; Cruces, 2005; Cruces and Wodon, 2007).

This article follows the VEP approach and defines vulnerability as the threat of future deprivation due to its intuitive interpretation and applicability. While the other approaches have desirable features, they often entail making more restrictive assumptions. For instance, the VEU approach requires imposing common utility and risk preferences (Just and Pope, 2003). Meanwhile, the VER approach unavoidably requires detailed and long running longitudinal data, which is mostly unavailable in Latin America.

Separating vulnerability and poverty

As stated in the introduction, vulnerability and poverty are two distinct but related phenomena. Accounting for their significant overlap and identifying them separately is a challenging task. The main motivation for this breakdown is policy-oriented, since the tools to alleviate poverty are not necessarily the same as those required to prevent it (Barrientos, 2007).

Until recently, the relationship between poverty and risk had been mostly unaccounted for in the distributive literature, which relies mostly on ex-post analysis, such as poverty assessments and profiles. While these provide meticulous cross-sectional views of deprivation, they fail to account for its dynamic characteristics. A series of recent studies have tried to fill this gap by developing forward-looking measures of vulnerability. Their basic premise is that households face different risks of either remaining or becoming poor. The distinction between vulnerability and deprivation is important because while all the poor are usually considered vulnerable, the converse is not necessarily true (Suryahadi *et al.*, 2000).

Studying vulnerability has a series of potential benefits. On the one hand, it helps identify household characteristics linked to future poverty. On the other hand, it also sheds light on coping mechanisms with regards to risk. Findings in these two dimensions could inform the policy design process and improve it. For instance, mechanisms which reduce vulnerability may be promoted (*e.g.*, better credit and insurance markets) and existing social safety nets may be strengthened to account for both idiosyncratic and aggregate risk.

The definition of vulnerability as the threat of deprivation is also related to recent efforts to classify the poor into those who are chronically (or structurally) poor and those who are transiently (or temporarily) poor. These studies have found that those who are observed to be always poor differ in their characteristics from the sizeable fraction of households experiencing temporary poverty, which is usually related to specific shocks (see, for instance, Jalan and Ravallion, 1998 and 2000, for China; and Cruces and Wodon, 2003, for Argentina). However there is an important conceptual and practical distinction between the two methods. While the transient/chronic poverty approach is ex-post or backward looking, the vulnerability literature attempts to capture the distribution of future welfare levels.

Vulnerability to poverty: The basic approach

The definition of vulnerability adopted in this document is the ex-ante risk that a household will be poor if it is currently not poor, or that it will remain in poverty if it is currently poor. This definition implies that vulnerability may best be summarized as a probability. Chaudhuri *et al.* (2002) and Chaudhuri (2003) represent this probability as:[4]

$$V_{ht} = \Pr(y_{h,t+1} \leq z) \tag{1}$$

where $y_{h,t+1}$ is a measure of household welfare at time $t+1$, and z is an exogenous predefined poverty line. To obtain vulnerability estimates, it is necessary to define the level of minimum acceptable welfare (the poverty line) and estimate the level of future welfare based on current data. The first element does not pose any significant issues. The second, however, is more complex. To estimate future welfare, it is necessary to make assumptions about how it is generated, which involves a discussion of its determinants and dynamics. As a starting point to address these concerns, consider a general reduced form for an income generating function:

$$y_{ht} = f(X_h, \beta_t, \alpha_h, e_{ht}) \tag{2}$$

where X_h represents a set of observable household and community characteristics, β_t is a vector of parameters at time t, α_h is an unobserved time-invariant household effect, and e_{ht} is a mean-zero disturbance term that captures idiosyncratic factors. Since the methodology obtains these estimates from a single point in time, the unobserved household level heterogeneity cannot be properly estimated. Nevertheless, this pitfall is overcome somewhat by including extensive information on household and community characteristics. Substituting Equation 2 into Equation 1, household vulnerability at time t may be rewritten as:

$$V_{ht} = \Pr\left(y_{h,t+1} = f(X_h, \beta_{t+1}, \alpha_h, e_{h,t+1}) \leq z | X_h, \beta_{t+1}, \alpha_h, e_{h,t+1}\right) \tag{3}$$

The above expression suggests that if proper estimates of future welfare may be obtained from cross-section data, then vulnerability may be feasibly estimated by Equation 3. Implicitly, this specification encompasses the fundamental identifying assumptions of the approach. Firstly, future levels of welfare are relatively stationary from one period to the next.[5] Secondly, welfare is determined by observable factors as well as unexpected shocks, *i.e.* poverty risk may be due either to low expected welfare or high volatility. This specification of the welfare-generating process, and thus its distribution, implies that both the mean and the variance need to be taken into account.

Therefore, the necessary steps to consistently estimate vulnerability using cross-sectional data are: 1) make distributional assumptions, 2) specify the welfare-generating process and estimate the relevant parameters from the data source, and 3) obtain the probability of being poor. The authors suggest that the ideal informational source to implement this method is panel data of sufficient length, since the availability of repeated observations adds a crucial dimension (variability) to measures of household welfare. However, given the scarcity of longitudinal data in developing countries, the authors argue that the validity of these techniques is also suitable when using cross-sectional information.

A review of recent vulnerability applications[6]

Chaudhuri *et al.* (2002) apply the above methodology to cross-sectional data from Indonesia. Their results demonstrate that the vulnerable population is generally larger than the fraction observed poor at a given point in time. The study also finds differences between the distribution of vulnerability and poverty across different population characteristics (*e.g.* regions, educational levels, *etc.*). Chaudhuri (2003) uses data from the Philippines and Indonesia with analogous results.

Other applications of vulnerability in cross-sectional settings reflect similar findings. For instance, Suryahadi and Sumarto (2003) analysed the effects of the 1997 economic crisis in Indonesia on vulnerability and found that measuring aggregate shocks is essential to identify those at risk properly. For Latin America, Tesliuc and Lindert (2004) study the case of Guatemala, while Gallardo (2009) concentrates on Nicaragua. In general, their evidence suggests that vulnerability is widespread, with vulnerable households usually outnumbering those who actually become poor. Moreover, these studies identify several household characteristics associated with vulnerability. These include household head characteristics, such as gender, education, employment status and area of residence.

Christiaensen and Subbarao (2005) extend this basic framework to estimate vulnerability to poverty using pseudo-panels, or a time series of cross-sections. Their application to rural Kenya indicates that idiosyncratic shocks substantially affect the volatility of consumption. The feasibility of creating these data sources motivated a number of ensuing studies.[7] Finally, a number of approaches to vulnerability measurement have employed panel data to obtain their estimates.[8] Studies using these data sources include Suryahadi *et al.* (2000), Kamanou and Morduch (2002), Chaudhuri (2003), McCulloch and Calandrino (2003).

This growing body of case studies and methodological developments on vulnerability has prompted a critical assessment of this framework.[9] Some studies, namely those that rely on panel data, undertake validation exercises of their cross-sectional vulnerability estimates by contrasting them with observed future individual poverty states and aggregate poverty rates (for instance, Chaudhuri *et al.*, 2002; Chaudhuri, 2003; and Zhang and Wan, 2009). The results of these exercises indicate that cross-sectional estimates of expected poverty seem to provide relatively good approximations of aggregate rates, although they do not test predictive power at household level.

Data and empirical strategy

Argentina: short panels (one year)

Given its rotating sampling structure from 1995 to 2003, Argentina's *Encuesta Permanente de Hogares* (EPH) enables the generation of panel data.[10] This structure implies that it is possible to track a fraction of the total sample for a period of time. In particular, 25 per cent of the sample could be tracked throughout four consecutive semesters. Or, 50 per cent could be potentially observed in one year intervals (see Cruces and Wodon, 2007). Attrition is not significant in the data, estimated at approximately 16 per cent of the sample (Gutiérrez, 2004) and seems to be random (Albornoz and Menéndez, 2007). This implies that any estimates from this source are unbiased.[11]

In this study, the data are assembled into yearly panels, *i.e.* the same household is observed once in the baseline and again one year later (during the month of October) using balanced panels, since attrition is not a significant source of bias.

In addition to observing households over a one year period, the rotating panel nature of the surveys implies that it is also possible to construct "cohorts" of households. The data allows for the assembling of a total of seven cohorts, from 1995-1996 until 2001-2002. The main advantage of this approach is that it captures behaviour during growth (1995-1998), recession (1999-2000) and crisis (2001-2002) episodes in Argentina. Thus, it provides a test of the vulnerability measure's sensitivity to changing macroeconomic conditions. The sample for Argentina is described in Panel A of Table 1.

Chile: long panels (five and ten years)

Chile's *Encuesta de Caracterización Socioeconómica* (CASEN) is the country's main socioeconomic survey. In 1996, the Statistics Institute selected 5210 households in four regions to be tracked over the coming years.[12] By 2009, two follow-up rounds were made available. The first corresponded to 2001, and the second to 2006. The main advantage of this longitudinal data is its span of ten years, which provides information on relatively long term outcomes.

The Chilean data allows tracking the same households throughout the entire timeframe, contrary to the Argentinean case where households are only followed for one year. Hence, an overall balanced panel would contain households observed in all three rounds (1996, 2001 and 2006). However, to test the predictive power of vulnerability estimates over both the medium and the long term, the analysis is carried out over three timeframes: two five-year panels (1996-2001 and 2001-2006), denoted as short term periods, and the long term period covering the initial and final rounds, 1996-2006. Another difference with the Argentine data is that attrition is higher in the CASEN panel, as Bendezú *et al.* (2007a) find. In fact, one quarter of the original sample dropped out in the first follow-up and by the last available survey only half the initial sample remained. Despite these problems, the potential bias is addressed using the longitudinal expansion factors provided with the data (see Bendezú *et al.*, 2007a). The sample for Chile is described in Panel B of Table 1.

Empirical strategy

In a previous section three steps were established to empirically estimate vulnerability from cross-section data. As a reminder these are: 1) distributional assumptions; 2) specification of the welfare generating process and estimating the relevant parameters from the data; and 3) obtaining the predicted probability of being poor.

Table 1
Descriptive statistics for the data and sample

Year	Households	Regions	Household Size	Children	Male Head	Years of education of Household Head
Panel A: Argentina						
1995-1996	9,174	5	3.9	1.3	77.1	8.7
1996-1997	8,712	5	3.8	1.2	74.9	8.9
1997-1998	7,392	6	3.8	1.2	74.9	8.9
1998-1999	8,012	6	3.8	1.2	73.0	9.0
1999-2000	7,170	6	3.8	1.2	73.0	9.1
2000-2001	7,053	6	3.7	1.2	72.0	9.3
2001-2002	6,829	6	3.8	1.1	71.0	9.1
Panel B: Chile						
1996-2001	3,090	4	4.2	1.3	74.9	8.0
2001-2006	3,090	4	4.0	1.0	71.6	8.4
1996-2006	3,090	4	4.1	1.2	71.5	8.4

Let us begin with the first step. In any distributive analysis, the selection of the welfare proxy is crucial for the resulting estimates. In this study, welfare is measured using household per capita income as surveys in Latin America do not regularly collect consumption or expenditure data. For the purposes at hand, the study follows usual convention and assumes that income is distributed as a lognormal random variable.[13] This approximation simplifies the estimation of vulnerability, since lognormal distributions can be fully characterized by their mean and variance.

Approaching the second step is less straightforward. Take the standard cross-sectional income model commonly used in applied work:

$$\ln y_h = X_h\beta + e_h \tag{4}$$

where X_h represents a set of observable household and community characteristics. In the estimates presented here, and based on previous work in vulnerability literature, the covariates in X_h include a series of structural characteristics of the household: the household head's gender and age (and age squared), household size and its square, number of young children in the household, number of employed members, head of household educational level (using educational categories), and whether the household is located in urban or rural areas. This general specification is selected primarily to increase comparability across the surveys and time, and constitutes a set of characteristics known to be related to structural poverty and the income generating process.[14] Finally, the error term e_h comprises all other unobservable effects.

However, due to the initial distributional assumption, the variance of expected income must also be estimated to compute the probability of future poverty. Chaudhuri *et al.* (2002) and Chaudhuri (2003) assume that the disturbance term e_h

captures community specific effects and idiosyncratic shocks on household income, and that its variance is correlated with observable household and environment characteristics. This explicitly assumes that the expected income variance is heteroscedastic. A simple parametric way to express this characteristic is to model the variance linearly:

$$\sigma^2_{e,h} = X_h \theta \tag{5}$$

As is known, standard regression analysis based on ordinary least squares (OLS) assumes homoscedasticity, making estimates of β and θ unbiased but inefficient if this assumption does not hold. To deal with this problem, Chaudhuri (2003) applies a three-step feasible generalized least squares (FGLS) method first proposed by Amemiya (1979) to obtain efficient estimates of these parameters. Using the consistent and asymptotically efficient estimators $\hat{\beta}$ and $\hat{\theta}$ obtained by FGLS, the expected log income and variance for each household may be obtained by calculating:

$$\hat{E}\left[\ln \hat{Y}_h | X_h\right] = X_h \hat{\beta}_{\text{FGLS}} \tag{6}$$

$$\hat{V}\left[\ln \hat{Y}_h | X_h\right] = \hat{\sigma}^2_{eh} = X_h \hat{\theta}_{\text{FGLS}} \tag{7}$$

Finally, Step 3 uses these estimates as inputs to compute the probability that a household will be poor in the future. Since income is assumed to be lognormal, the estimated conditional probability may be obtained by:

$$\hat{V}_h = \hat{\Pr}\left(\ln \hat{y}_{h,t+1} \leq \ln z | X_h\right) = \Psi\left(\frac{\ln z - X_h \hat{\beta}}{\sqrt{X_h \hat{\theta}}}\right) \tag{8}$$

where Ψ denotes the cumulative density of the standard normal distribution and z is defined as the 4 USD international poverty line expressed in 2005 purchasing power parity (PPP) terms (see Ravallion *et al.*, 2009). The selection of the 4 USD line responds to its growing use in the distributive literature for Latin America, mostly due to its similarity to the moderate official poverty lines of many countries, as is the case in both Argentina and Chile (see Gasparini *et al.*, 2012).

Some additional issues related to the estimation of income variance arise in the procedure outlined above. Firstly, there may be systematic measurement error in the observed welfare outcome. Income has a tendency to be underreported in household surveys, which may lead to underestimation of its variance and consequently bias vulnerability estimates upward. One solution involves scaling up the variance to account for this measurement error. However, given that the measurement error generating process is unknown, this study makes no adjustments to avoid imposing further assumptions. Therefore, if measurement error implies an underestimation of income variance, the estimates presented here may be regarded as a lower-bound of the probability of future poverty. Secondly, the linear specification of the variance implies that there might be negative estimates of the variance for certain households. If this proportion of households is high, then vulnerability estimates may be affected. However, in practice this problem is found to be minor (less than 1 per cent of the total sample), and thus negative observations were dropped.

Defining the state of vulnerability

The probabilities obtained from Equation 8 may be presented, interpreted and discussed in several ways.[15] Since the objective of the validation exercise is to quantify the performance of vulnerability estimates as predictors of future poverty, households are classified into categories: vulnerable and not vulnerable. This implies setting a probability threshold above which households are considered to be at risk for future deprivation. In general, there seems to be a consensus in the applied literature in using two thresholds: the current poverty rate, and a value of 0.50.[16] These two indicators will be referred to as the *relative* and *absolute* vulnerability thresholds, respectively. In the majority of estimates (see the two coming sections), the study uses the absolute threshold to define vulnerable states; the relative threshold is employed in the comparative final analysis of this article.

Vulnerability measures as predictors of future poverty

This section presents a series of validation exercises for the cross-sectional vulnerability framework, which consists of comparing predicted levels of vulnerability with future realised welfare outcomes, much in the spirit of time series one-step-ahead forecasts. Specifically, cross-sectional vulnerability estimates at time t are compared to realised outcomes in $t+1$. The evaluation proceeds in two stages.

The first stage computes mean vulnerability (or the average probability of future poverty) for the entire sample at time t, and compares it to the observed poverty rate in $t+1$. This estimate provides insight on whether vulnerability captures current and future aggregate poverty levels, and the magnitude of any potential discrepancies.

The second stage is more elaborate. The analysis focuses on misclassifications with respect to the entire population. In this part of the exercise, the proportion of households incorrectly classified is calculated using the total population as a reference point. This allows estimating the overall error at household level as the sum of those households, which were classified as vulnerable but did not become poor, and the non-vulnerable households, which actually became poor. The results are presented in matrix form shown in Table 2.

Global misclassifications can be computed as $M = (b+c)/N$. It should be stressed that even if the income generating process is correctly specified and cross-sectional data provides enough information for an assessment of each household's

Table 2
Definition of misclassified households

t	t+1 Poor	Non-poor	TOTAL
Expected poor	a Correctly classified	b Misclassified	**EP**
Expected non-poor	c Misclassified	d Correctly classified	**ENP**
TOTAL	**P**	**NP**	**N**

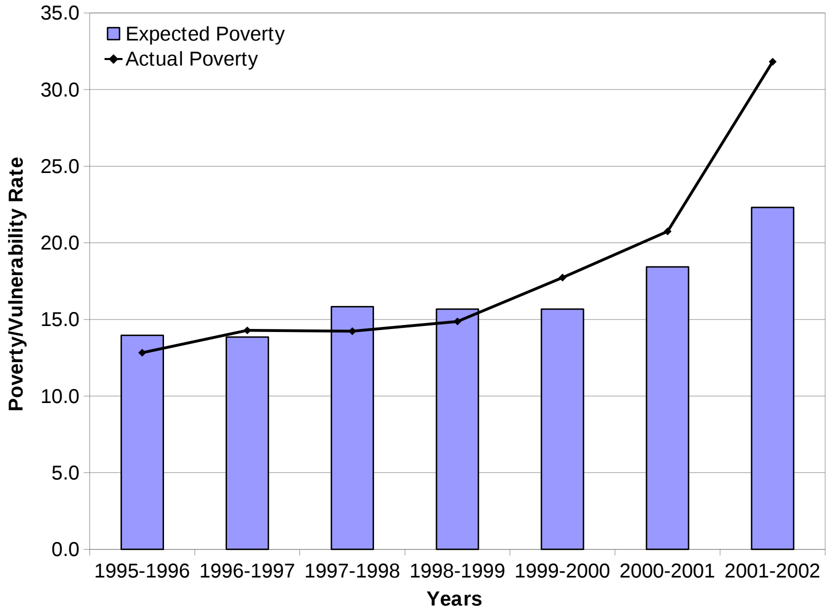
Source: Authors' calculations on Argentina and Chile panel data.

probability of becoming poor in the following period, one should not expect all vulnerable households to be poor and all non-vulnerable to be non-poor in $t+1$ since this is a probabilistic and not an exact prediction. However, this extreme case provides a plausible metric for quantifying errors.

On the other hand, misclassifications can also be computed with respect to a more restricted reference population — for instance, the poor. This is an important distinction because if the proportion of poor households is relatively small, misclassifications might appear high with respect to this group, but low with respect to the total population. The intuition for the relevance of these classification errors is best exemplified by a potential policy application.

For a policymaker devising a transfer-based safety net, vulnerable households (those with high probabilities of becoming poor in the future) constitute the target beneficiaries. In this scenario, misclassifying vulnerable households as non-vulnerable carries a high exclusion cost. These households would not receive the transfer, despite the fact that they would require it. In keeping with the statistics literature, this type of misclassification can be labelled as Type I (exclusion) error, corresponding to the proportion of currently poor households, which were classified as not vulnerable in the previous period. In the previous notation, this case would correspond to Type I: b/p. The second type of misclassification implies labelling non-vulnerable households as vulnerable — those more likely on average to become poor, but did not. In this case, these households would not require the transfer. These inclusion errors can be labelled as Type II, and correspond to the fraction of currently non-poor households, which were classified as vulnerable in the preceding period, or Type II: c/NP. From the policymaker's perspective, weighting equity over efficiency, Type I errors seem more serious than Type II errors, although budgetary concerns might change this perspective.

Figure 1
Argentina: Vulnerability as expected poverty and actual poverty



Source: Authors' calculations on Argentina panel data.

The predictive power of vulnerability measures: Short run evidence from Argentina

The results of the aggregate validation for Argentina are presented in Figure 1, which plots the expected poverty rate computed from the information available in t and the actual poverty rate in the second year of each cohort $(t+1)$ to observe errors at the aggregate level.[17]

In general, with the exception of the last cohort, which covers the extraordinary macroeconomic crisis of 2001-2002, expected poverty levels kept fairly close to actual poverty rates. The divergence increases from the 1999-2000 cohort onward, which coincides with the start of the recession that culminated in the crisis. At the onset of the recession, vulnerability underestimated actual poverty. This problem was exacerbated during the 2001-2002 crisis, when the vulnerability assessment based on 2001 data grossly underestimated the 2002 poverty rate by more than 10 percentage points.[18] This substantial underestimation highlights the difficulties of accounting for exogenous future shocks in a cross-sectional setting.[19]

Hence, the validation exercise indicates that vulnerability estimates in the short run predict aggregate poverty relatively well during periods of stability, when the stationarity assumption is more likely to hold. However, in the case of negative shocks, there is a clear risk of underestimating future poverty. This finding implies that the presence of a positive shock may lead to overestimating poverty. In an extreme case, the difference may be quite substantial. However, these shocks must be particularly strong (as during the 2001-2002 crisis in Argentina) to cause significant deviations. Therefore, these estimates may be considered as lower bounds for future poverty in the absence of external shocks.

The results for the micro-level validations are presented in Table 3 for each cohort of the Argentinean panels. The results for M (the overall misclassification indicator) demonstrate that 86 per cent of all households are classified correctly (averaging results for all cohorts). This total corresponds to 79 per cent of non-poor households and 7 per cent of poor households. The remaining 14 per cent of households are classified incorrectly, with 3 per cent corresponding to non-poor households in $t+1$ deemed vulnerable in t, and 11 to poor households in $t+1$ classified as not vulnerable with data from period t.

Keeping the same indicators for each cohort, there is clear evidence of a higher precision of the estimates during growth and stability periods, when almost 90 per cent of all cases are correctly classified. Entering the recession (the 1999-2000 cohort), M drops to 85 per cent. The worst rate is found in 2001-2002, when precision of predicted poverty falls by more than 10 percentage points to 79 per cent. Once again, it becomes clear that vulnerability estimates are sensitive to unaccounted shocks, resulting in increased error levels. However, it should be stressed that more than three quarters of total households are correctly classified, although these figures refer to proportions over the whole population. Classification errors with respect to those in poverty show a different picture.

Estimates of Type I and Type II errors are presented in Figure 2. These estimates may be interpreted as the percentage of incorrectly classified households with respect to the entire poor population (Type I) and the non-poor population (Type II). The results in these tables indicate that, on average, more than 61 per cent of the poor are wrongfully classified. The fraction of Type II (inclusion) errors is substantially lower, ranging from 3 to 4 per cent for all cohorts. During growth periods, Type I error is greater (64 per cent in 1995), actually improving slightly during recession (63 per cent in 1999) and the crisis (58 per cent in 2001). However, this improvement is small in magnitude. Even in the best case scenario, more than half of those who become poor are not classified as vulnerable using this method. The opposite is true for Type II (inclusion) errors; in worse aggregate economic conditions, the amount of non-poor classified as vulnerable increases. Nevertheless, the results indicate that the magnitude of the imprecision is small.

In general, the findings for short term panels from Argentina suggest that although estimates of vulnerability classify a substantial majority of all households correctly, misclassification errors are substantial when focusing only on the poor. In fact, 3 out of 5 poor households would be categorized as not vulnerable.[20] These findings cast doubts on the usefulness of cross-sectional vulnerability estimates for

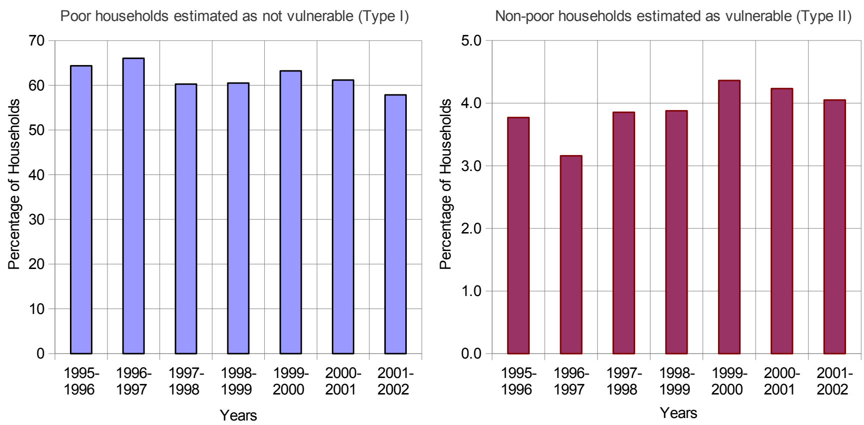
Table 3
Argentina: Misclassifications

	1996	
1995	**Poor**	**Non-poor**
Expected poor	4.6	3.3
Expected non-poor	8.3	83.9
	1997	
1996	**Poor**	**Non-poor**
Expected poor	4.9	2.7
Expected non-poor	9.4	83.0
	1998	
1997	**Poor**	**Non-poor**
Expected poor	5.7	3.3
Expected non-poor	8.6	82.5
	1999	
1998	**Poor**	**Non-poor**
Expected poor	5.9	3.3
Expected non-poor	9.0	81.8
	2000	
1999	**Poor**	**Non-poor**
Expected poor	6.5	3.6
Expected non-poor	11.2	78.7
	2001	
2000	**Poor**	**Non-poor**
Expected poor	8.1	3.4
Expected non-poor	12.7	75.9
	2001	
2002	**Poor**	**Non-poor**
Expected poor	13.4	2.8
Expected non-poor	18.4	65.4

Source: Own calculations on Argentina panel data.
Note: All calculations use as the denominator the entire population.

targeting aid programs at household level. Additionally, the evidence also shows that the effect of aggregate shocks on Type I and II errors is relatively minor. In this case study, Type I misclassifications remain at a high level and Type II misclassifications are always low, regardless of the overall state of the economy.

Figure 2
Argentina: Evolution of misclassified households. Estimated Type I and Type II errors.



Source: Authors' calculations on Argentina panel data.
Notes:
(1) Type I households are the fraction of poor households in $t+1$ which are classified as not vulnerable in t.
(2) Type II households are the fraction of non-poor households in $t+1$ which are classified as vulnerable in t.

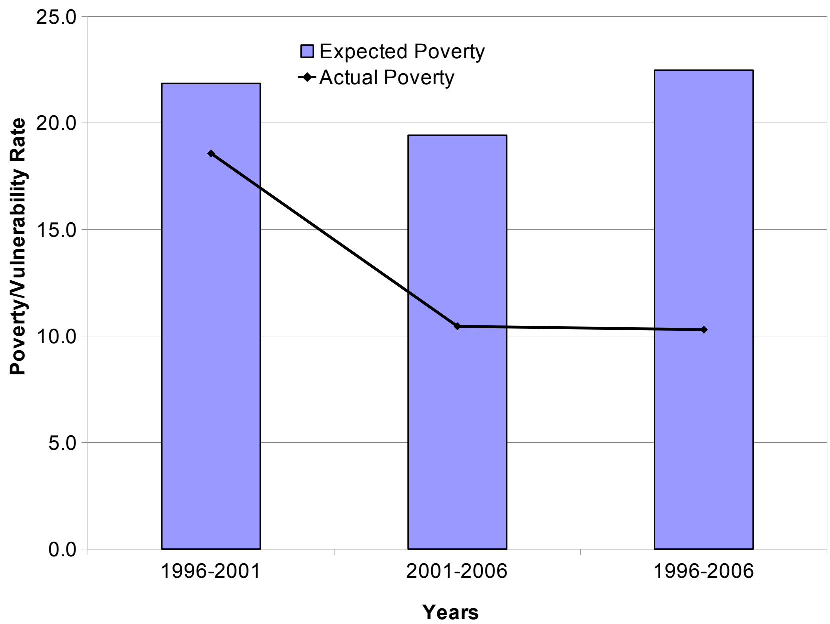
The predictive power of vulnerability measures: Long run evidence from Chile

In this section, the same evaluation is carried out based on data with a substantially longer timeframe. During this period the Chilean economy did not experience the large aggregate fluctuations observed in the Argentinean case; but rather a sustained period of growth (between 4 and 6 per cent per year) and poverty reduction.

Although the longer timeframe setting suggests a lesser degree of income persistence than with yearly data,[21] it is still possible that the variables capturing a household's income-generating process are better suited to predicting long term prospects rather than short term fluctuations. Whether vulnerability estimates fare better over longer periods is ultimately an empirical question, which the following estimates seek to clarify.

Estimates for the aggregate validation are depicted graphically in Figure 3. In general, the results indicate that vulnerability overestimates actual poverty in Chile, contrary to the results for Argentina. This suggests that during sustained periods of growth the method is more 'pessimistic', since its design cannot account for diminishing poverty trends. Moreover, this feature seems to be exacerbated by the length of the timeframe considered. For instance, for both short term periods the difference in expected and realised poverty is between 3 to 9 percentage points, and 12 for the longest period. This evidence indicates that cross-sectional vulnerability estimates are less precise in predicting future poverty in the long run, at least where the presence of marked trends in poverty is concerned.

Figure 3
Chile: Vulnerability as expected poverty and actual poverty



Source: Authors' calculations on Chile panel data.

The results for the micro-level validation exercise of the Chilean case are presented in Table 4. The level of misclassification, M, indicates that 84 per cent of total households are classified correctly when averaging all time periods. The remaining 16 per cent of households are classified incorrectly, with 9 per cent corresponding to non-poor households deemed vulnerable, and 7 per cent to poor households classified as not vulnerable. The magnitude of these results remains unchanged when the analysis focuses on short or long term periods.

Figure 4 summarizes calculations for Type I and Type II errors for Chile, and plots the evolution of both error types for each period. This figure reveals that, on average, more than half of the poor are incorrectly classified by the method. It is noteworthy that although this type of error is relatively high (especially from a targeting perspective), its magnitude is lower in comparison to the estimates for Argentina. Also, the fraction of Type II (inclusion) errors ranges between 9 to 12 per cent for all time periods, which is more than three times that for Argentina. Comparing both types of errors, the results show that in the long run, the method

Table 4
Chile: Misclassifications

	2001	
1996	**Poor**	**Non-poor**
Expected poor	9.2	7.3
Expected non-poor	9.5	74.0
	2006	
2001	**Poor**	**Non-poor**
Expected poor	3.9	8.6
Expected non-poor	6.3	81.1
	2006	
1996	**Poor**	**Non-poor**
Expected poor	4.7	11.5
Expected non-poor	5.4	78.4

Source: Own calculations on Chile panel data.
Note: All calculations use as the denominator the entire population.

performs just as ineffectively when focusing on poor households, but that it also falters with respect to the non-poor.

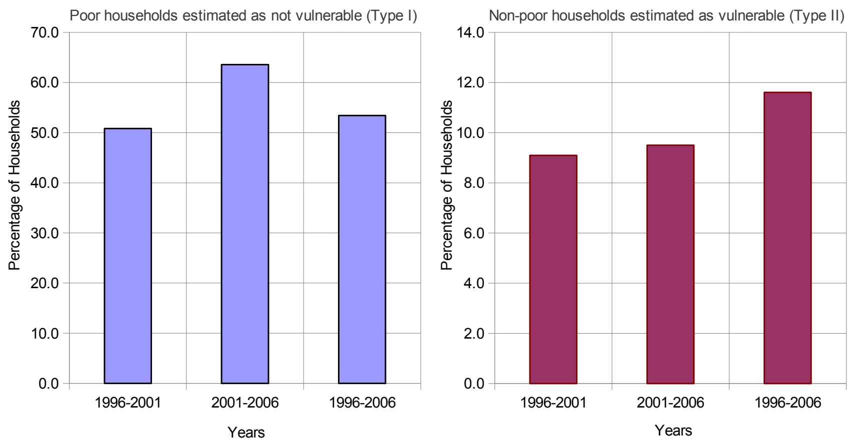
In general, the findings for long run panel data confirm that the cross-sectional vulnerability estimates classify most households correctly when taking the entire population as a reference point. However, when focus is placed on the poor, the level of misclassification is high, with the method classifying roughly half of poor households incorrectly. The validation exercise reveals that cross-sectional vulnerability estimates do not perform noticeably better or worse over a longer period.

The predictive power of vulnerability across the income distribution

The results of these validation exercises indicate that vulnerability estimates have a relatively high degree of misclassification, especially among the poor. However, it should be noted that these classification errors are average estimates, which can mask heterogeneities across the income distribution. Since vulnerability measures are mainly motivated as tools to capture welfare variability among those below and close to the poverty line, this section analyzes the issue of misclassification across income groups.

The decomposition exercise presented below estimates Type I and Type II errors by income deciles. The income deciles are specified at time t, when vulnerability is estimated, and the errors are defined in $t+1$.[22] As in the previous section, these validation exercises rely on panel data, and mimic policymakers' problems in assigning limited resources in $t+1$ based on information collected in t, and using

Figure 4
Chile: Evolution of misclassified households. Estimated Type I and Type II errors.



Source: Authors' calculations on Chile panel data.
Notes:
(1) Type I households are the fraction of poor households in $t+1$ which are classified as not vulnerable in t.
(2) Type II households are the fraction of non-poor households in $t+1$ which are classified as vulnerable in t.

the realised status in $t+1$ to measure the indicator's effectiveness.

The general structure of the results presented in the tables below is as follows: the "fraction poor in $t+1$" column presents the participation of each decile in the relevant population, *i.e.* for Type I (or Type II) errors, the proportion of poor (or non-poor) households as a function of the decile they occupied during the previous period. These proportions offer an *ad hoc* indicator of mobility as they indicate from where in the distribution in t the poor in $t+1$ come from. The following column summarizes group-specific errors. The average error presented in the prior sections may be obtained as a weighted average of these errors (using the proportions in the first column as weights).

Tables 5 and 6 present the decomposition results for Argentina. The results in Table 5 indicate that most of the future poor are located in the lower-end of the original income distribution, particularly among the first three deciles. Within these groups, the vulnerability measure is most effective for households in the first decile, with values of the exclusion error (Type I) around 36-38 per cent, and with a very low value of 13.8 per cent corresponding to the 2001-2002 crisis. Exclusion errors are substantially higher for the next two deciles. Finally, the very high exclusion errors (above 90 per cent) for households above the median of income distribution represent, in fact, a relative methodological success. As indicated by the "fraction

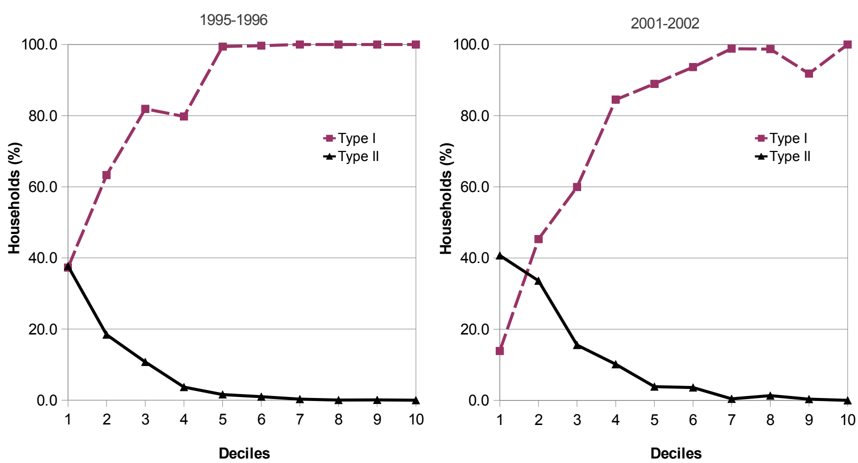
Table 5
Argentina: Type I (exclusion) errors by income decile

Decile in t	1995-1996		1996-1997		1997-1998		1998-1999	
	Fraction poor in $t+1$	Type I Error	Fraction poor in $t+1$	Type I Error	Fraction poor in $t+1$	Type I Error	Fraction poor in $t+1$	Type I Error
1	0.309	37.3	0.308	36.9	0.414	43.1	0.377	36.4
2	0.322	63.3	0.336	66.1	0.300	61.6	0.277	55.9
3	0.168	81.9	0.145	88.1	0.115	75.2	0.133	88.2
4	0.069	79.8	0.069	96.9	0.101	82.3	0.081	89.0
5	0.058	99.4	0.064	89.8	0.026	99.6	0.052	86.5
6	0.036	99.7	0.026	100.0	0.023	98.6	0.028	94.1
7	0.003	100.0	0.013	96.6	0.006	100.0	0.034	100.0
8	0.016	100.0	0.022	76.0	0.007	100.0	0.009	100.0
9	0.020	100.0	0.012	100.0	0.007	100.0	0.006	100.0
10	0.000	100.0	0.005	100.0	0.001	100.0	0.004	100.0
Overall error		**64.3**		**66.0**		**60.3**		**60.5**

Decile in t	1999-2000		2000-2001		2001-2002	
	Fraction poor in $t+1$	Type I Error	Fraction poor in $t+1$	Type I Error	Fraction poor in $t+1$	Type I Error
1	0.346	38.4	0.309	38.3	0.185	13.8
2	0.273	61.9	0.302	53.5	0.239	45.3
3	0.156	77.7	0.160	77.7	0.242	59.9
4	0.094	87.9	0.102	86.0	0.135	84.5
5	0.070	95.0	0.063	89.7	0.084	89.0
6	0.024	97.0	0.034	97.7	0.051	93.7
7	0.019	100.0	0.015	100.0	0.041	98.8
8	0.014	100.0	0.012	100.0	0.014	98.7
9	0.002	100.0	0.002	100.0	0.008	91.8
10	0.001	64.8	0.001	100.0	0.001	100.0
Overall error		**63.2**		**61.2**		**57.8**

poor in $t+1$" column, there are very few better-off households that end up poor in the following period. The methodology classifies most of these households as non-vulnerable because of their income generating capacity in t and cannot be expected to capture these few outliers. These general findings hold irrespective of aggregate economic conditions, as shown in Figure 5, which compares a relatively stable period (1995-1996) with a deep aggregate crisis (2001-2002). The results in Table 6 indicate that Type II errors are highest in the poorest deciles, although these represent a small proportion of the future non-poor (on average, less than 3 per cent of the non-poor in $t+1$ were in the first decile in t). This confirms that vulnerability estimates are relatively effective at predicting who will be non-poor across the entire distribution.

Figure 5
Argentina: Errors by income decile



Source: Authors' calculations on Argentina panel data.
Notes:
(1) Deciles are defined at time t.
(2) Type I error is the fraction of poor households in $t+1$ which were classified as not vulnerable in t.
(3) Type II error is the fraction of non-poor households in $t+1$ which were classified as vulnerable in t.

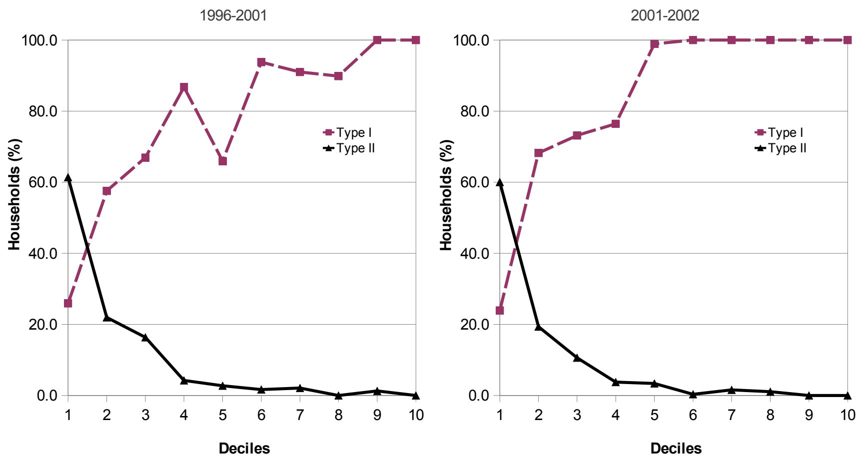
Tables 7 and 8 present the same results for the long panels. As with the yearly data, a large fraction of the poor in $t+1$ (2001) or $t+2$ (2006) were located in the first three deciles of the per capita income distribution in the initial period t (1996). The vulnerability estimates have substantially lower exclusion errors for the lowest decile. Table 8 indicates that inclusion errors are highest among the poor, who represent a small fraction of the future non-poor. The predictions are quite precise for the middle and upper end of the income distribution (see Figure 6).

Table 6
Argentina: Type II (inclusion) errors by income decile

Decile in t	1995-1996 Fraction poor in $t+1$	1995-1996 Type II Error	1996-1997 Fraction poor in $t+1$	1996-1997 Type II Error	1997-1998 Fraction poor in $t+1$	1997-1998 Type II Error	1998-1999 Fraction poor in $t+1$	1998-1999 Type II Error
1	0.024	37.7	0.024	46.2	0.032	32.5	0.030	40.3
2	0.062	18.4	0.063	12.5	0.070	17.9	0.068	17.1
3	0.089	10.7	0.101	5.2	0.098	7.1	0.084	9.5
4	0.109	3.7	0.105	3.3	0.103	4.1	0.112	3.8
5	0.105	1.6	0.110	2.4	0.119	2.7	0.109	1.7
6	0.125	1.0	0.112	0.8	0.109	0.8	0.117	0.7
7	0.111	0.3	0.121	0.2	0.119	0.2	0.118	0.1
8	0.125	0.0	0.118	0.1	0.120	0.0	0.125	0.1
9	0.118	0.1	0.120	0.0	0.113	0.0	0.122	0.1
10	0.130	0.0	0.123	0.0	0.115	0.0	0.116	0.0
Overall error		**3.8**		**3.2**		**3.9**		**3.9**

Decile in t	1999-2000 Fraction poor in $t+1$	1999-2000 Type II Error	2000-2001 Fraction poor in $t+1$	2000-2001 Type II Error	2001-2002 Fraction poor in $t+1$	2001-2002 Type II Error
1	0.018	34.1	0.009	42.1	0.014	40.7
2	0.063	22.0	0.041	27.5	0.020	33.6
3	0.096	12.7	0.085	15.2	0.048	15.5
4	0.101	6.1	0.096	7.2	0.090	10.1
5	0.116	1.1	0.112	3.8	0.111	3.8
6	0.121	1.3	0.119	1.8	0.117	3.6
7	0.120	1.0	0.133	0.5	0.120	0.4
8	0.124	1.1	0.134	0.1	0.164	1.3
9	0.122	0.0	0.135	0.0	0.152	0.3
10	0.120	0.0	0.136	0.0	0.166	0.0
Overall error		**4.4**		**4.2**		**4.1**

Figure 6
Chile: Errors by income decile



Source: Authors' calculations on Chile panel data.
Notes:
(1) Deciles are defined at time t. (2) Type I error is the fraction of poor households in $t+1$ which were classified as not vulnerable in t.
(3) Type II error is the fraction of poor households in $t+1$ which were classified as vulnerable in t.

Table 7
Chile: Type I (exclusion) errors by income decile

	1996-2001		2001-2006		1996-2006	
Decile in t	**Fraction poor in $t+1$**	**Type I Error**	**Fraction poor in $t+1$**	**Type I Error**	**Fraction poor in $t+1$**	**Type I Error**
1	0.429	25.9	0.416	35.7	0.492	23.9
2	0.235	57.5	0.204	73.7	0.149	68.2
3	0.132	66.9	0.177	81.6	0.069	73.2
4	0.096	86.8	0.066	93.7	0.106	76.5
5	0.038	66.0	0.060	98.8	0.070	98.9
6	0.030	93.8	0.032	86.1	0.013	100.0
7	0.024	91.0	0.018	92.5	0.004	100.0
8	0.008	89.9	0.005	100.0	0.016	100.0
9	0.002	100.0	0.001	100.0	0.004	100.0
10	0.006	100.0	0.022	100.0	0.077	100.0
Overall error		**50.8**		**63.6**		**53.4**

It is natural to question whether and how the proportion of vulnerable households may be affected by the choice of threshold. If this is so, its impact may influence the results presented above.[23] As a way to examine the sensitivity of the assessment to the selected cut-off points, the following exercise calculates the

Table 8
Chile: Type II (inclusion) errors by income decile

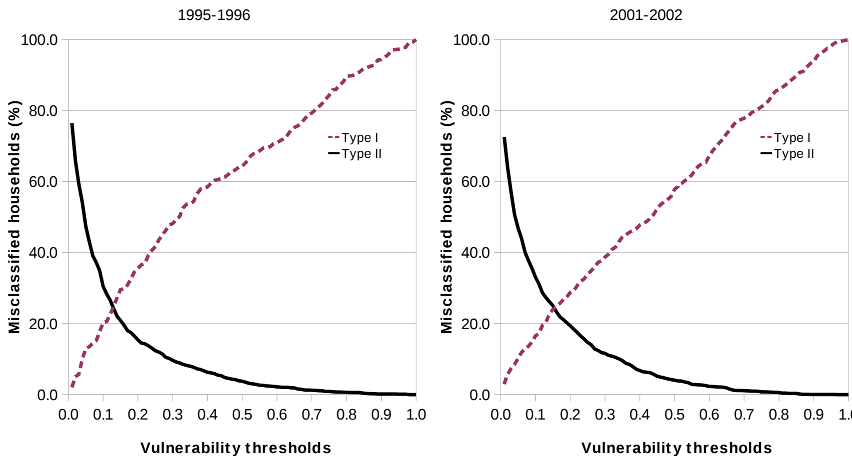
Decile in t	1996-2001		2001-2006		1996-2006	
	Fraction poor in $t+1$	Type II Error	Fraction poor in $t+1$	Type II Error	Fraction poor in $t+1$	Type II Error
1	0.066	61.4	0.072	50.0	0.109	60.0
2	0.097	22.0	0.125	28.2	0.143	19.3
3	0.099	16.3	0.096	15.4	0.124	10.6
4	0.118	4.2	0.107	5.9	0.115	3.8
5	0.105	2.7	0.113	1.6	0.084	3.4
6	0.093	1.6	0.088	0.8	0.097	0.3
7	0.111	2.1	0.121	0.1	0.085	1.6
8	0.097	0.0	0.090	0.0	0.082	1.1
9	0.108	1.3	0.086	0.0	0.083	0.0
10	0.106	0.0	0.101	0.0	0.076	0.0
Overall error		**9.1**		**9.5**		**11.6**

percentage of Type I and Type II misclassifications for *all* possible vulnerability thresholds for both the Argentinean and Chilean data. Results are shown for two specific periods: 1995-1996 (growth) and 2001-2002 (crisis) in case of Argentina, and 1996-2001 and 2001-2006 for Chile, respectively.

Figures 7 and 8 plot the percentage of Type I and Type II misclassifications on the vertical axis, with the corresponding thresholds in the horizontal axis. The results indicate that both types of errors respond differently as the threshold increases. For instance, Type I error increases markedly with the cut-off, which could be expected. Intuitively, as the threshold rises, more households are considered to be vulnerable, and in the extreme case the error should be equal to the poverty rate. Also as expected, Type II errors fall as the threshold rises. The largest possible error of this type is attained when the vulnerability threshold is set at its lowest value: all households would be considered as vulnerable and the error would be equal to the proportion of non-poor households.

Where exactly should researchers set the vulnerability threshold to minimize these errors? An intuitive procedure is to select the point where both lines intersect, which corresponds to the threshold that minimizes the sum of both types of error, implicitly assuming equal weights. In the estimates for Argentina, it seems that this 'optimum' value of the threshold is close to 0.15, while it is roughly 0.20 for Chile. Interestingly, for both countries these figures are close to the observed poverty rates in both periods (see Figures 1 and 3, respectively), with the exception of the Argentine crisis period in which the poverty rate increased to 30 per cent, well above its trend. Overall, these findings would seem to indicate that a relative threshold, *i.e.* the poverty line, is a relatively good rule of thumb to minimize classification errors, at least for those samples.

Figure 7
Argentina: Estimated Type I and Type II errors for all possible vulnerability thresholds



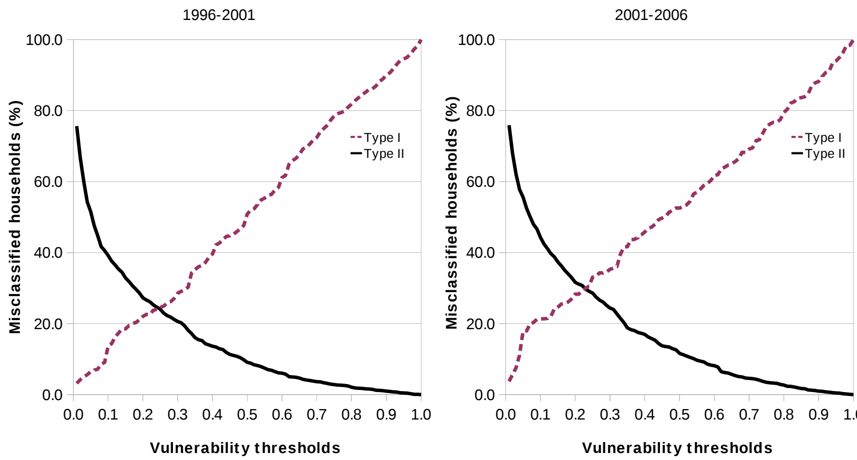
Source: Authors' calculations on Argentina panel data.
Notes:
(1) Type I households are the fraction of poor households in $t+1$ which are classified as not vulnerable in t.
(2) Type II households are the fraction of non-poor households in $t+1$ which are classified as vulnerable in t.

Nonetheless, it is essential to remember that these results apply to countries with medium levels of poverty (in an international comparison), where the absolute and relative thresholds differ substantially. Ultimately, the choice of threshold depends on the proposed objective of the vulnerability estimates. If the purpose is to target social programs, then the vulnerability threshold should perhaps be chosen endogenously, in a manner which minimizes some weighted average of the Type I and Type II errors. These weights should reflect the preference and judgment of the researcher or policymakers.

A comparative assessment of deprivation indicators

The above evaluation indicates that cross-sectional vulnerability estimates seem to misclassify many households, although this error is lower for those at the bottom of income distribution. This result, however, lacks a benchmark for comparison. This section carries out a comparative assessment of several deprivation indicators' capacity to identify the future poor as a means to determine whether vulnerability significantly adds to a policymaker's toolbox. It thus provides the possibility of contrasting the performance of the vulnerability measure relative to other indica-

Figure 8
Chile: Estimated Type I and Type II errors for all possible vulnerability thresholds



Source: Authors' calculations on Argentina panel data.
Notes:
(1) Type I households are the fraction of poor households in $t+1$ which are classified as not vulnerable in t.
(2) Type II households are the fraction of non-poor households in $t+1$ which are classified as vulnerable in t.

tors. The deprivation measures discussed below include alternative specifications of vulnerability (using absolute and relative thresholds), regression-based income predictions, indicators of basic needs deficits and multidimensional poverty measures.[24] Even though some of these deprivation indicators (for instance, multidimensional poverty measures) were not designed with the purpose of predicting future risk or outcomes, which is an explicit objective of vulnerability measures, their application in policy settings (*e.g.* in Mexico, Honduras and Nicaragua) justifies their inclusion in this exercise.

The strategy in this section consists of three main steps. First, the analysis computes each deprivation measure for each household in period t, and classifies the population in terms of broadly defined vulnerability groups (they are classified as vulnerable if deprived according to the indicator and not vulnerable otherwise). The second step compares this classification with observed poverty in $t+1$, which allows obtaining Type I and Type II errors for each indicator. Finally, these errors are presented for the two poorest deciles of income distribution to measure predictive power for households with the lowest income.

Figures 9-12 graphically present estimates of exclusion and inclusion errors for the first and second deciles of the income distribution (defined in period t) for the selected indicators and periods. The first two bars on the left of each figure correspond to vulnerability estimates using the absolute threshold, which will be taken as the point of comparison (see Appendix). The following bars summarize results for vulnerability using the relative cut-off (poverty rate), income predictions, unsatisfied basic needs (UBN), and different specifications of the Alkire-Foster multidimensional deprivation measure (A&F).

For Argentina's one year panels, the results for Type I (exclusion) errors indicate a relatively wide range in the performance of the indicators for households in the first decile (first panel of Figure 9). The results are qualitatively similar for the second decile of income distribution (second panel of Figure 9), but the level of exclusion errors increases for all measures, indicating more efficiency in identifying the chronic poor.

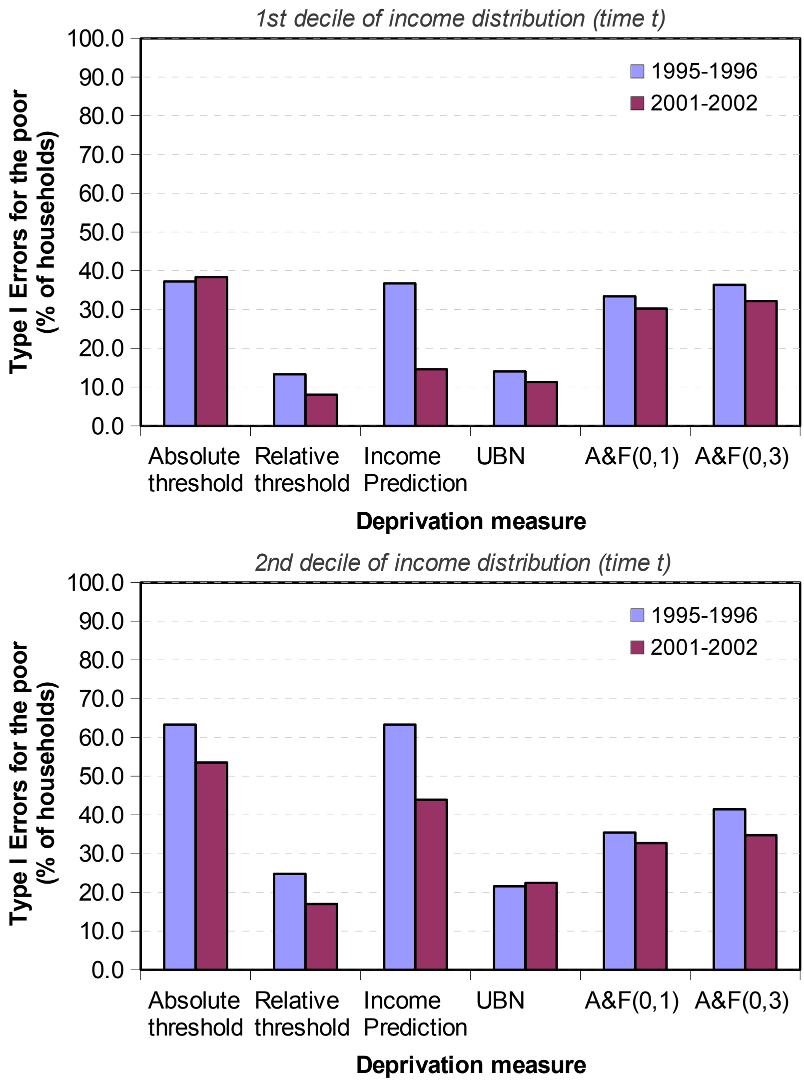
In general, the measure with the highest level of accuracy in terms of Type I error is vulnerability using the relative threshold (10 per cent error on average), followed closely by UBN. The worst performers for the two poorest deciles are vulnerability with the absolute threshold, income predictions and the A&F measures. These conclusions seem robust regardless of the aggregate conditions. In contrast, Type II (inclusion) errors demonstrate the opposite behaviour. The cases of the UBN and the vulnerability measure based on a relative threshold, which perform well in terms of low exclusion errors, reveal relatively high Type II errors (Figure 10).

This behaviour highlights the observed trade-off between the two types of error since, as described above; minimizing exclusion errors leads to larger inclusion errors. Ultimately, the budget assigned to social programs and the costs of information gathering will lead to a cost-benefit analysis, and will determine where the line is to be drawn for these conflicting errors (see Ravallion and Chao, 1989, for a more detailed discussion of targeting trade-offs).

The results for the longer term Chilean panels are similar to those for Argentina (Figure 11). For the first decile of per capita income, the vulnerability measure based on a relative threshold is fairly accurate in identifying the future poor, revealing levels of exclusion error at less than 8 per cent in both selected time periods. However, unlike for Argentina, vulnerability with the absolute threshold (the standard or most commonly used measure of vulnerability) appears to be more effective, with low error levels close to the results from UBN measures. With regards to the second decile of income distribution, the magnitude of exclusion errors increases for all indicators, while the relative measure of vulnerability appears, yet again, to be the most effective. For Type II (inclusion) errors, the same trade-off between inclusion and exclusion is evident (Figure 12). Vulnerability based on a relative threshold and UBN demonstrate the highest levels of inclusion errors regardless of the decile or time span.

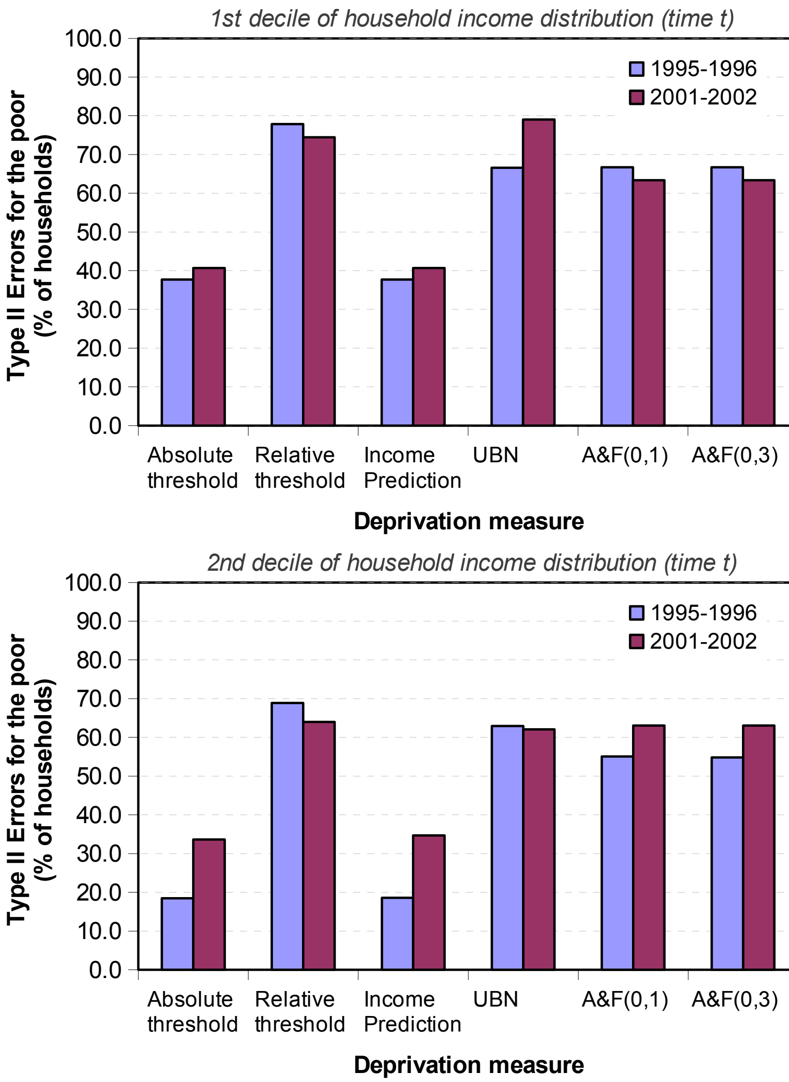
Figure 9

Argentina: Type I (exclusion) errors for selected deprivation measures



Source: Authors' calculations on Argentina panel data.
Notes:
(1) A household is considered poor if its expected log household income (obtained by Equation 4) is below the log poverty line.
(2) The basic needs considered are: number of rooms in house, house location, house materials, water, restroom, children's education, education of household head and number of earners. A household is considered as poor if they meet at least one of the above conditions.
(3) Multidimensional A&F$(0,k)$ refers to the dimension-adjusted headcount ratio proposed by Albire and Foster (2011). The parameter k is the cut-off across dimensions. The dimensions considered are income, education, overcrowding, access to water and housing quality.

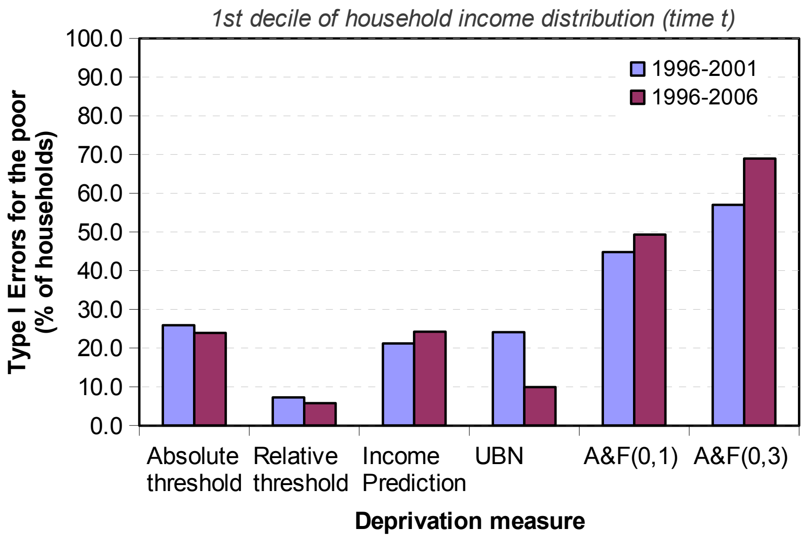
Figure 10
Argentina: Type II (inclusion) errors for selected deprivation measures

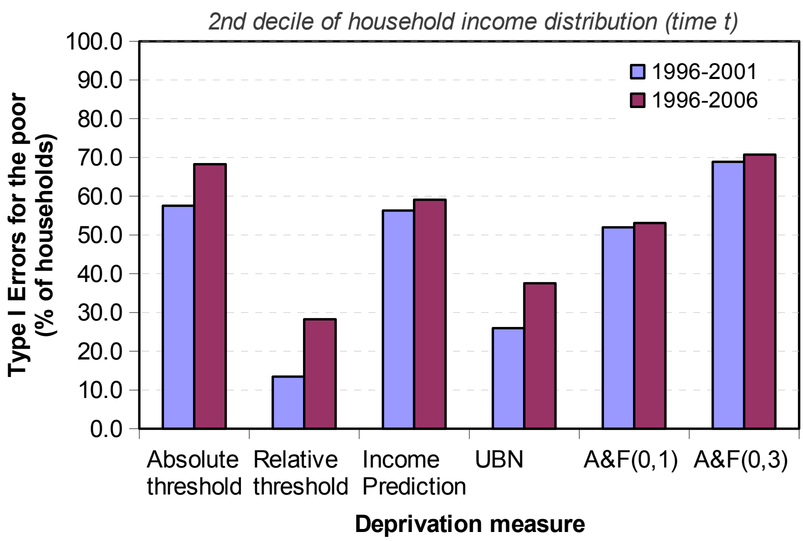


Source: Authors' calculations on Argentina panel data.
Notes:
(1) A household is considered poor if its expected log household income (obtained by Equation 4) is below the log poverty line.
(2) The basic needs considered are: number of rooms in house, house location, house materials, water, restroom, children's education, education of household head and number of earners. A household is considered as poor if they meet at least one of the above conditions.
(3) Multidimensional A&F$(0,k)$ refers to the dimension-adjusted headcount ratio proposed by Albire and Foster (2011). The parameter k is the cut-off across dimensions. The dimensions considered are income, education, overcrowding, access to water and housing quality.

Figure 11

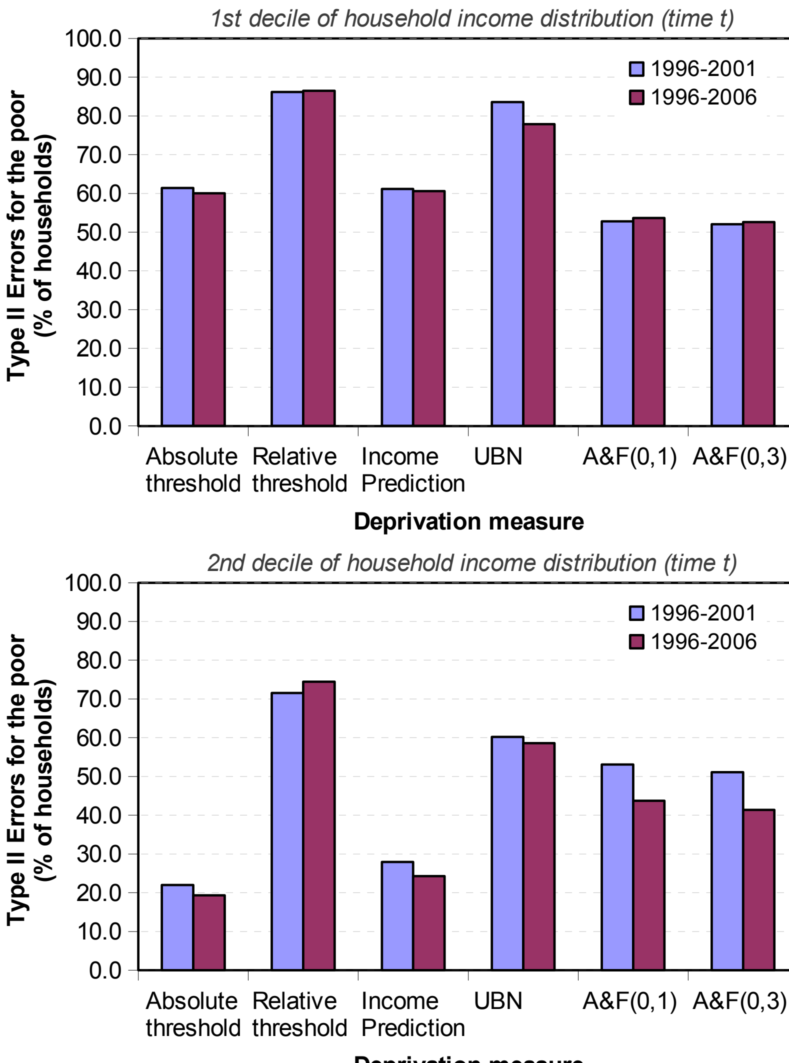
Chile: Type I (exclusion) errors for selected deprivation measures





Source: Authors' calculations on Argentina panel data.
Notes:
(1) A household is considered poor if its expected log household income (obtained by Equation 4) is below the log poverty line.
(2) The basic needs considered are: number of rooms in house, house location, house materials, water, restroom, children's education, education of household head and number of earners. A household is considered as poor if they meet at least one of the above conditions.
(3) Multidimensional A&F$(0,k)$ refers to the dimension-adjusted headcount ratio proposed by Albire and Foster (2011). The parameter k is the cut-off across dimensions. The dimensions considered are income, education, overcrowding, access to water and housing quality.

Figure 12
Chile: Type II (inclusion) errors for selected deprivation measures



Source: Authors' calculations on Argentina panel data.
Notes:
(1) A household is considered poor if its expected log household income (obtained by Equation 4) is below the log poverty line.
(2) The basic needs considered are: number of rooms in house, house location, house materials, water, restroom, children's education, education of household head and number of earners. A household is considered as poor if they meet at least one of the above conditions.
(3) Multidimensional A&F$(0,k)$ refers to the dimension-adjusted headcount ratio proposed by Albire and Foster (2011). The parameter k is the cut-off across dimensions. The dimensions considered are income, education, overcrowding, access to water and housing quality.

Conclusions

This study assessed the effectiveness of cross-section vulnerability measures, defined as their predictive power of future poverty states at the aggregate and household level using short term (Argentina) and long term (Chile) panel data. In both cases, the findings suggest that measures of vulnerability classify most households correctly when taking the entire population as a reference point, while demonstrating a relatively high level of misclassification when the poverty status of individual households over time is concerned. Errors are, however, substantially lower among households in the bottom 10 and 20 per cent of the income distribution. The specific contexts of both case studies (wide aggregate fluctuations for Argentina, sustained growth and falling poverty for Chile) illustrate both the possibilities and limitations of cross-sectional estimates of vulnerability as predictors of future poverty.

The validation exercise also compared the predictive power of vulnerability measures with respect to deprivation indicators. The comparative assessment indicated that the lowest exclusion errors are attained with vulnerability measures based on a relative threshold and UBN indicators at the cost of high inclusion errors.

These results suggest that cross-sectional vulnerability estimates might provide useful information for analysts and policy makers, but that the results from the methodology need to be complemented with further information. For instance, vulnerability profiles should help to distinguish which poor households classified as not vulnerable are truly experiencing a temporary poverty spell, and which ones are true classification errors. Moreover, the estimates can benefit greatly from information on overall economic conditions, or on aggregate or group-specific shocks. At the same time they can inform policymakers of distributional trends without full national household surveys (Mathiassen, 2009). While the exercises presented here analysed the performance with respect to monetary income, assessing the effectiveness of different deprivations using a wider set of dimensions is an interesting direction for future research.

Appendix: Alternative deprivation indicators

Inability to generate income

Haveman and Bershadker (1998), among other authors, have focused their analysis of poverty on the ability of households to generate resources, rather than on their effective availability. They define a household's capacity to generate income as the sum of the potential earnings of its members, based on observable characteristics. The authors attempt to structurally model income generation capacity, use the results from the regressions to compute fitted values for income, and compare this potential income with an exogenous poverty line. The analysis presented here uses fitted values from Equation 4 to obtain household income predictions. These values classify households as vulnerable if the fitted values of income are below the poverty line in $t+1$, and as not vulnerable otherwise. The difference between this methodology and the one used throughout this study is that vulnerability measures include a further transformation of the predicted income as it implies computing the conditional probability of being poor. The comparison of the Haveman and Bershadker (1998) approach and the vulnerability measures provides a benchmark to test whether this additional step adds information or mitigates measurement error over the simple income prediction.

Unsatisfied Basic Needs (UBN)

The Unsatisfied Basic Needs (UBN) approach is a non-income method widely used in Latin America (most notably by ECLAC, see Santos *et al.*, 2010) to capture structural poverty at the household level. The approach classifies a household as poor according to the UBN criterion if it exhibits a deficit in at least one the following dimensions (see Santos *et al.*, 2010, for specific details of the dimensions employed here):

- Overcrowding: more than 4 dwellers per room
- The household's dwelling is located in a 'poor or precarious' location (*e.g.* shanty towns)
- The dwelling is made of low-quality materials
- The dwelling does not have access to a water network
- The dwelling does not have a hygienic restroom
- There are children aged 7 to 11 not attending school
- The household head does not have a primary school degree
- High dependency ratio: a combination of two conditions, the household head does not have a high school degree and there are more than 4 household members for each income earner.

Deprivation as UBN is a 'union' indicator. Hence households are classified as vulnerable if they have deficiencies in at least one of the above dimensions and not vulnerable otherwise.

Multidimensional deprivation

This section also estimates one measure of the family of multidimensional poverty indicators developed by Alkire and Foster (2011). The criterion identifies the poor in two stages, first by defining a threshold for each considered dimension; and second, by exogenously defining the number of dimensions in which the household should be deprived to be considered poor. The second stage allows evaluating both union (poor in at least one dimension) and intersection (poor in all dimensions) criteria, but is flexible enough to allow for intermediate cases. Once identified, the poor are aggregated by a counting approach based on the Foster, Greer and Thorbecke — FGT — (1984) measures of poverty.

Specifically, the analysis below employs the dimension-adjusted headcount ratio measure (hereafter, $\text{A\&F}(0,k)$) which is the result of two components: a multidimensional headcount ratio (H); and the average deprivation share across the poor (A). Formally, it is defined as:

$$\text{A\&F}(0,k) = HA = \frac{1}{nd}\sum_{i=1}^{n} c_i \pi_k(x_i;z) \tag{A1}$$

where d represents the number of considered dimensions, n the number of households in the sample population, x_i is the outcome of household i in dimension k and z the deprivation line for that dimension. c_i depicts the sum of weighted deprivations for each household.[25] The term $\pi_k(x_i;z)$ represents a multidimensional identification function relating to a cut-off level k, such that it takes value 1 if $c_i \geq k$, indicating that the household is multidimensionally poor (taking value 0 if otherwise). The aggregation of $\pi_k(x_i;z)$ across the sample population results in the number of poor q_k, identified by both sets of cut-offs. Taking averages, this provides the multidimensional headcount ratio H. On the other hand, A is obtained by summing the (weighted) deprivations of all poor households and dividing by the maximum number of possible deprivations. In words, A represents the fraction of possible dimensions d in which the average multidimensionally poor household is deprived.

Therefore, $\text{A\&F}(0,k)$ can be expressed as a product between the percentage of multidimensional poor (H) and the average deprivation share across the poor (A). It may thus be interpreted as a headcount measure adjusted by the fraction of (weighted) dimensions in which poor households are deprived. The advantage of the weighting adjustment is that it allows the measure to satisfy a desirable property, monotonicity across dimensions. The $\text{A\&F}(0,k)$ measure estimated here uses the dimensions and thresholds in Table A1.

All dimensions are equally weighted, which assumes a "neutral" criterion about each component's relative importance. The inclusion of both "structural" and money metrics of poverty follows the criteria set by Battistón *et al.*'s (2009) application to

Table A1
Definition of dimensions and thresholds for A&F$(0,k)$ measure

Dimension	Indicator	Weight	Threshold
Education	Education of household head in years	1	6 years
Income	Per capita income	1	U$S 4 poverty line
Overcrowding	Persons per room	1	3 persons per room
Acces to water	Dwelling has access to water	1	Yes/No
Housing quality	Dwelling is made of low-quality materials	1	Yes/No

Latin America. Finally, deprived households are defined in three ways: $k = 1,2,5$. The first corresponds to a union approach, the second to an intermediate case, and the last to the intersection approach.

Notes

[1]*Acknowledgements:* The authors wish to thank Armando Barrientos for encouraging this work, the CPRC reviewers who provided valuable comments to earlier versions of this document, the participants at the 2010 CPRC Conference "10 Years of 'War against Poverty' ", as well as two anonymous referees for helpful comments and suggestions. All remaining errors are the sole responsibility of the authors.

[2]Marcelo Bérgolo and Guillermo Cruces are currently affiliated with the Center for Social, Labor and Distributional Studies (CEDLAS), Universidad Nacional de La Plata, and the National Scientific and Technical Research Council (CONICET), Argentina. At the time of writing, Ham was also affiliated to these institutions. Guillermo Cruces is also affiliated to IZA (Institute for the Study of Labor, Bonn, Germany).

[3]Mainly, the most important issue highlighted by this literature is that the VEP approach is shortsighted since it focuses on income and thus omits other relevant dimensions of welfare.

[4]This is the simplest approach to vulnerability measurement. A series of extensions are possible, since the vulnerability indicators are based on the Foster *et al.* (1984) measures of poverty. In the general Chaudhuri (2003) setup, this case corresponds to the headcount index (FGT with $\alpha = 0$).

[5]Of course, this is a highly restrictive assumption, which also depends on the timeframe from which conclusions are meant to be drawn. For instance, it is probable that welfare does not vary significantly from one year to the next. However, as this extrapolation period expands, the assumption is not likely to hold. See Christiaensen and Subbarao (2005) for a discussion of this topic.

[6]For a more complete literature review, which includes the VEU and VER approaches to vulnerability, see Hoddinott and Quisumbing (2008).

[7]For instance, see Bourguignon *et al.* (2004), Naudé *et al.* (2009) and Zhang and Wan (2009).

[8]Further details about the benefits and limitations of panel data are discussed below.

[9]Some studies highlight the limitations of measuring the probability of becoming poor and suggest using the expected squared poverty gap (Christiaensen and Subbarao, 2005; and Kamanou and Morduch, 2002). Another common problem involves determining the time horizon to assess vulnerability. For instance, Suryahadi *et al.* (2000) define household vulnerability as the probability of observing at least one spell of poverty in n periods, instead of only one (usually the following period).

[10]The survey design changed in 2003 from the rotating panels to a continuous sampling framework.

[11]It should be emphasized that problems of attrition and measurement errors may influence poverty estimates and estimates of other relevant variables in studies based on panel data (Alderman *et al.* 2001; Baulch and Hoddinott, 2000). Alderman *et al.* (2001) analyze the extent and implications of attrition for three developing countries and conclude that attrition can bias the estimates of several outcomes and certain family background variables. Their findings, however, suggest that attrition does not generally affect the consistency of coefficient estimates in linear regressions and models with categorical dependent variables. The methodology in this article relies on linear regressions models.

[12]The households belong to the third, seventh, eighth and metropolitan regions.

[13]See McDonald (2008) for more on this matter.

[14]In the results presented here, Equation 4 is estimated separately by time-comparable geographic regions within each country. This disaggregated estimation strategy accounts for potential differences in the structure

of local economies, a source of heterogeneity, which would be unaccounted for when estimating the equations at national level. For details about the regions and their definitions, see SEDLAC (CEDLAS and World Bank, 2010).

[15]See the working paper, Cruces *et al.* (2010) for results using additional vulnerability indicators, different poverty lines and vulnerability thresholds. The main findings presented here are robust to changes in these specifications.

[16]However, the discussion on vulnerability thresholds is open to debate since it shares many of the common characteristics with poverty lines (see Reddy and Pogge, 2010). In this case, since the exercise aims to test the predictive power of vulnerability, the thresholds are taken as given.

[17]The point estimates for this Figure and the following ones are available in the tables presented in Cruces *et al.* (2010).

[18]As noted by Chaudhuri *et al.* (2002), vulnerability estimates will probably differ from future poverty rates in the presence of large shocks, but with no group-specific shocks, average expected poverty should coincide with the current (rather than the future) poverty rate. This effect is apparent in Figure 1. While not necessarily an accurate predictor of future poverty, expected poverty is fairly similar to the same year's observed value.

[19]The stationarity assumption in the income equation plays a fundamental role here given that this decision does not contemplate potential shocks to the economy that might have a direct impact on welfare outcomes. However, modelling these shocks into the economy is not straightforward. Moreover, most Latin American economies do not depend strongly on observable shocks (*e.g.*, rainfall and other factors related to climate). Without data on these factors, their inclusion does not seem to be feasible. See Ferreira *et al.*, (2004), and Hoddinott and Quisumbing (2008) for a discussion of how to incorporate shocks into income equations.

[20]While this is considered a misclassification according to the benchmark defined in the previous setting, some of households might in fact be experiencing a temporary poverty spell, while having structural characteristics that make them non-vulnerable. This possibility is discussed below.

[21]Correlations for the five-year periods are lower than those estimated for Argentina by almost 20 percentage points (ranging between 0.52-0.62).

[22]Using income groupings in $t+1$ would be counterintuitive. For instance, exclusion (Type I) errors are defined over households who become poor in $t+1$. Hence, income deciles in this period would concentrate only on the lower-end of the distribution, omitting movements across the entire spectrum. The logic is the same for inclusion errors, although in this case the reference population is the non-poor.

[23]The authors are grateful to an anonymous referee for suggesting this exercise.

[24]The specification of these indicators is detailed in the Appendix.

[25]Each dimension has a specific weight. The weights are such that they add up to the total number of dimensions d.

Bibliography

Albornoz, F. and Menendez, M. (2007) "Income dynamics in Argentina during the 1990s: 'Mobiles' did change over time", *Económica* LIII, 1–2: 21–51.

Alderman, H., Behrman, J., Kohler, H., Maluccio, J., and Watkins, S. (2001) "Attrition in Longitudinal Household Survey Data", *Demographic Research* 5, 4: 79–124.

Alkire, S. and Foster, J. (2011) "Counting and Multidimensional Poverty Measurement", *Journal of Public Economics* 95, 7–8: 476–487.

Amemiya, T. (1977) "The Maximum Likelihood estimator and the non-linear three stage least squares estimator in the general nonlinear simultaneous equation model", *Econometrica* 45: 955–968.

Barrientos, A. (2007) "Does vulnerability create poverty traps?", CPRC Working Paper, No. 76, Chronic Poverty Research Centre.

Battistón, D., Cruces, G., López Calva, L.F., Lugo, M., and Santos, M. (2009) "Income and Beyond: Multidimensional Poverty in six Latin American countries", Oxford Poverty & Human Development Initiative (OPHI) Working Paper, No. 17, and CEDLAS Working Paper, No. 90.

Baulch, B. and Hoddinott, J. (2000) "Economic Mobility and Poverty Dynamics in Developing Countries", *Journal of Development Studies* 36, 6: 1–24.

Bendezú L., Denis, A., and Zubizarreta, J.R. (2007a) "Análisis de la atrición de la muestra de la Encuesta Panel CASEN", Documento de Trabajo No. 1, Observatorio Social, Universidad Alberto Hurtado.

Bourguignon, F., Goh, C. and Kim, D.I. (2004) "Estimating individual vulnerability to poverty with pseudo-panel data", World Bank Policy Research Working Paper, No. 3375, The World Bank, Washington D.C.

CEDLAS and The World Bank (2010) *SEDLAC Socio-Economic Database for Latin America and the Caribbean*.

Available at: `http://sedlac.econo.unlp.edu.ar/`

Chaudhuri, S., Jalan, J., and Suryahadi, A. (2002) “Assessing household vulnerability to poverty from cross-sectional data: A methodology and estimates from Indonesia”, Economics Department Discussion Papers, No. 117, Columbia University, New York.

——— (2003) “Assessing vulnerability to poverty: Concepts, empirical methods and illustrative examples”, Mimeo, Columbia University, New York.

Christiaensen, L.J. and Subbarao, K. (2005) “Towards an Understanding of Household Vulnerability in Kenya”, *Journal of African Economies* 14, 4: 520–558.

Cruces, G. (2005) “Income Fluctuations, Poverty and Well-Being Over Time: Theory and Application to Argentina”, STICERD-DARP Working Paper, No. 76, London School of Economics, London.

Cruces, G., and Wodon, Q.T. (2003) “Transient and chronic poverty in turbulent times: Argentina 1995-2002”, *Economics Bulletin* 9, 3: 1–12.

——— (2007) “Risk-Adjusted Poverty in Argentina: Measurement and Determinants”, *Journal of Development Studies* 43, 7: 1189–1214.

Cruces, G., Gasparini, L., Bérgolo, M., and Ham, A. (2010) “Vulnerability to Poverty in Latin America: Evidence from Cross-Section and Panel Data”, CPRC Working Paper, No. 170, Chronic Poverty Research Centre.

Dercon, S. (2006) “Vulnerability: A microperspective”, in Bourguignon, F., Pleskovic, B., and van der Gaag, J. (eds.), *Securing development in an unstable world*, The World Bank, Washington D.C.

Ferreira, F., Leite, P.G., Pereira da Silva, L.A., and Pichetti, P. (2004) “Can the Distribution impact of Macroeconomic Shocks be predicted? A comparison of the performance of macro-micro models with historical data for Brazil”, World Bank Policy Research Working Paper, No. 3303, The World Bank, Washington D.C.

Foster, J.E., Greer, J., and Thorbecke, E. (1984) “A Class of Decomposable Poverty Measures”, *Econometrica* 52, 3: 761–766.

Gallardo Altamirano, M. (2009) “Estimación de corte transversal de la vulnerabilidad y la pobreza potencial de los hogares en Nicaragua”, CEDLAS Working Paper, No. 86, Universidad Nacional de la Plata, Argentina.

Gasparini, L., Cicowiez, M., and Sosa Escudero, W. (2012) *Pobreza y Desigualdad en América Latina: Conceptos, herramientas y aplicaciones*, Buenos Aires: Editorial Temas. In press.

Gutierrez, F. (2004) “Dinamica Salarial y Ocupacional: Analisis de panel para Argentina 1998-2002”, CEDLAS Working Paper, No. 11, Universidad Nacional de la Plata, Argentina.

Haveman, R and Bershadker, A. (1998) “Self Reliance as Poverty Criterion: Trends in Earnings-Capacity Poverty, 1975-1992”, *The American Economic Review* 88, 2: 342–347.

Hoddinnott, J. and Quisumbing, A.R. (2008) “Methods for Microeconometric Risk and Vulnerability”, Mimeo, International Food Policy Research Institute, Washington D.C.

Jalan, J. and Ravallion, M. (1998) “Transient Poverty in Postreform Rural China”, *Journal of Comparative Economics* 26, 2: 338–357.

——— (2000) “Is Transient Poverty Different? Evidence for Rural China”, *Journal of Development Studies* 36, 6: 82–99.

Just, R.E. and Pope, R.D. (2003) “Agricultural Risk Analysis: Adequacy of Models, Data and Issues”, *American Journal of Agricultural Economics* 85, 5: 1249–1256.

Kamanou, G. and Morduch, J. (2002) “Measuring Vulnerability to Poverty”, in Dercon, S. (ed.), *Insurance against poverty, UNU/WIDER Studies in Development Economics*, Oxford University Press.

Ligon, E., and Schecter, L. (2003) “Measuring vulnerability”, *The Economic Journal* 113, 486: C95–C102.

Mathiassen, A. (2009) “A model based approach for predicting annual poverty rates without expenditure data”, *Journal of Economic Inequality* 7, 2: 117–135.

McCulloch, N. and Calandrino, M. (2003) “Vulnerability and chronic poverty in rural Sichuan”, *World Development* 31, 3: 611–628.

McDonald, J.B. (2008) “Some Generalized Functions for the Size Distribution of Income”, Chotikapanich, D. (ed.), *Modeling income distributions and Lorenz curves*, Springer, Verlag.

Naudé, W., Santos-Paulino, A.U., and McGillivray, M. (2009) “Measuring Vulnerability: An Overview and Introduction”, *Oxford Development Studies* 37, 3: 183–191.

Ravallion, M. and Chao, K. (1989) “Targeted Policies for Poverty Alleviation under Imperfect Information: Algorithms and Applications”, *Journal of Policy Modeling* II, 2: 213–224.

Ravallion, M., Chen, S., and Sangraula, P. (2009) “Dollar a day revisited”, *The World Bank Economic Review* 23, 2: 163–184.

Reddy, S. and Pogge, R. (2010) “How Not To Count the Poor”, in Anand, S., Segal, P., and Stiglitz, J. (eds.), *Debates on the Measurement of Global Poverty*, Oxford University Press.

Santos, M., Lugo, M., López Calva, L.F., Cruces, G., and Battistón, D. (2010) “Refining the Basic Needs Approach: A multidimensional analysis of poverty in Latin America”, *Research on Economic Inequality* 18.

Suryahadi, A. and Sumarto, S. (2003) “Poverty and Vulnerability in Indonesia before and after the Economic Crisis”, *Asian Economic Journal* 17, 1: 45–64.

Suryahadi, A., Sumarto, S., and Pritchett, L. (2000) “Quantifying vulnerability to poverty: A proposed measure applied to Indonesia”, World Bank Policy Research Working Paper, No. 2437, The World Bank, Washington D.C.

Tesliuc, E., and Lindert, K. (2004) “Risk and vulnerability in Guatemala: A quantitative and qualitative assessment”, Social Protection Discussion Paper, No. 30154, The World Bank, Washington D.C.

Zhang, Y. and Wan, G. (2009) “How precisely can we estimate vulnerability to poverty?”, *Oxford Development Studies* 37, 3: 277–287.