Measuring poverty persistence with missing data

With an application to Peruvian panel data

Yadira Diaz Cuervo

Stephen Pudney

ISER, University of Essex

This version: October 22, 2013

Abstract

We consider the estimation of measures of persistent poverty in panel surveys with missing data, focusing on the persistent poverty headcount, its duration-adjusted variant, and a related measure used by the European Union as an indicator of the risk of persistent poverty. We develop a partial identification approach to allow for data missing-not-at-random, and apply it to panel data from Peru for 2007-11. The "worst case" bounds are very wide, but we achieve much more precise identification by adding a set of weak a priori restrictions. Standard non-response weighting adjustments cannot be relied upon to remove missing-data bias.

Keywords: Missing data; Poverty persistence, Partial identification

JEL codes: D31, I32

Contact: Steve Pudney, ISER, University of Essex, Wivenhoe Park, Colchester, CO4 3SQ, UK; tel. +44(0)1206-873789; email spudney@essex.ac.uk

We are extremely grateful to Lucia Gaslac and her team at the Peruvian National Statistics Office who provided us with the datasets to make this work possible. Pudney's contribution to the research was supported by the Economic and Social Research Council through the MiSoC Research Centre (award no. RES518285001) We are grateful to participants in the Banca d'Italia economics seminar for valuable comments.

1 Introduction

Poverty is always a serious issue for the people directly affected by it, and for society in general. But the adverse long-term consequences of poverty are particularly serious for a society with highly persistent poverty concentrated in part of the population. Poverty persistence interferes with the processes of health and human capital formation, particularly early in life, and this propagates poverty across generations. Persistent poverty is also socially divisive. If we accept the importance of poverty persistence as a social problem, it becomes important to monitor its occurrence and change over time. A leading example of this policy concern for poverty persistence in developed countries is the adoption by the European Union (EU) of a specific measure of the risk of persistent poverty as one of its core social indicators: see Atkinson et al (2002) and European Commission (2009) for an account of the EU approach to monitoring of social exclusion. But poverty persistence is a much more challenging problem in developing countries, and is the focus of much research and policy activity (Baulch and Hoddinott 2000).

Poverty persistence is an ambiguous concept, and there are many possible measures. In discrete time, such measures capture the tendency for periods of poverty to cluster in adjacent periods for households observed longitudinally. Jenkins and Van Kerm (2012) have criticised the EU measure on grounds that single-period poverty rates provide a good indication of the evolution of persistent poverty over time and are less subject to the measurement difficulties that are associated with measurement of persistence. Nevertheless, poverty persistence remains an important concern, particularly in the development context, where poverty processes may be quite different to those observed in the EU.

Empirical implementation of any chosen measure of poverty persistence is not straightforward. It requires longitudinal household income or consumption data, which are inevitably affected by problems of non-response and sample attrition. Methods such as inverse probability weighting for non-response are widely used to adjust for these problems, but they rest on strong assumptions which are not directly testable. Typically, those assumptions require that the response process should be independent of the poverty process, conditional on observable characteristics and circumstances.

An alternative to pursuing the bias correction approach is to make fewer assumptions and accept that measurement will be subject to a degree of inherent uncertainty. This is the partial identification approach, which generates estimates of logical bounds on the persistence measure, rather than a single point estimate. There is a long history of this approach in statistics, but the current interest in partial identification largely stems from the pioneering work of Manski and his collaborators (Manski 1995, 2003). See Tamer (2010) for a recent review of the field. These methods have been used to deal with the problem of missing income data in poverty measurement in the single-period cross-section context by Nicoletti (2010) and Nicoletti et al (2011) but, to the best of our knowledge, they have not so far been used for poverty analysis in the dynamic context.

In this paper, we consider measures of poverty persistence which are either defined as the proportion of panel members experiencing at least some critical number of periods of poverty within a fixed observation window, or a simple elaboration of it. We focus on the problem of non-response/attrition and establish worst-case nonparametric bounds on the persistent poverty measures, showing that they are very wide for the rate of non-response typical of household panel surveys. This demonstrates that standard balanced-panel measurement of persistent poverty is highly vulnerable to doubts about the validity of weighting methods designed to compensate for non-response. We consider how the bounds can be narrowed by introducing additional information in the form of plausible a priori assumptions, and apply these improved bounds to the ENAHO panel data from Peru. We find that the resulting estimates are informative and that they cast serious doubt on the validity of standard weighting procedures conventionally used to 'adjust' for non-response.

2 Persistent poverty in Peru

In this paper, we analyse poverty at the household, rather than individual, level. There are several reasons for this. First, the household is the basic unit of anti-poverty policy in Peru (and many other countries), since the provision of assistance to particular individuals (especially children) within households is often infeasible. Poverty research at the household level matches this aspect of policy. Second, the conventional measures of economic wellbeing (such as equivalised income or consumption expenditure) that are used in most poverty analysis are defined at the household level. This implies that the multivariate distribution of individual poverty states with households is statistically degenerate in the sense that, with probability 1, all household members are in the same poverty state. Thus any individual-level poverty analysis is essentially a household analysis, weighted by household size. There is a third difficulty with individual-level poverty analysis in our cases, where we are dealing with missing data. If a household is non-respondent to the survey, we cannot observe the number of individuals it contains, so an individual-based analysis is impossible.

2.1 The Peruvian ENAHO panel surveys

Our analysis is based on the Peruvian National Household Survey, *Encuesta Nacional de Hogares* (ENAHO), which is the official household survey run by the Peruvian National Institute of Statistics since 1995 to monitor household living conditions. The information provided by ENAHO is used by the government to calculate official estimates of poverty, under the supervision of an independent poverty committee. The survey covers the whole national territory for urban and rural areas, 24 counties (departmentos) and the constitutional province of Callao. The sample is drawn from a sampling frame based on the 2005 National population and housing census, independently across counties. Within counties, a probabilistic sample of areas is drawn using a multistage, stratified procedure. For urban areas, primary sampling units are population centers with 2000 or more inhabitants and secondary sampling units are clusters with 120 addresses on average; tertiary sampling units are addresses. In rural areas, primary sampling units are of two types, either urban population centres with 500 to 2000 inhabitants, or rural enumeration areas covering an average of 100 addresses. The secondary sampling units for rural areas are either clusters of addresses (120 per cluster on average) or individual addresses. Tertiary sampling units for rural areas are addresses. On average there are six sampled addresses per urban cluster and eight per rural cluster. To measure changes in the behavior of some characteristics of the population, the ENAHO maintains a sub-sample of addresses as a panel. We use the five-wave panel spanning the period 2007 to 2011.

Like many panel surveys from developing countries, the ENAHO sampling scheme is address-based rather than household-based. Households that move to a new address are not followed and retained in the panel; instead the new occupants of the sampled address enter the panel as a separate new household unit. ENAHO conventions define a move to have taken place between period t and $t+1$ if there is no individual living at the address in both periods. An address affected by a (single) move will generate two panel households, one with missing observations for the periods before the move, the other with missing observations after the move. The "missingness" process therefore encompasses both the usual survey refusal and non-contact processes and the process of geographical mobility. The volume of missing data and the resulting bias consequently differs substantially from what would we would find in a household panel survey that follows mobile individuals according to a set of following rules. As has been pointed out by several authors (see Thomas et al 2001, Rosenzweig 2003 and Dercon and Shapiro 2007), mobility – particularly from rural to urban areas – can be an important way out of poverty for many households in developing countries, so there is potential for serious attrition bias and our approach, which avoids the missing at random (MAR) assumption underlying typical survey weighting schemes, is particularly valuable for such surveys.

The ENAHO collects a wide range of information on income and expenditure, allowing the construction of household annual aggregates of either total income or total expenditure. We use total annual household expenditure per equivalent adult rather than income as our welfare indicator for three reasons: first, expenditure is generally considered a more reliable indicator than income, which is often poorly measured in poor rural households; second, expenditure smoothes short-term fluctuations in household resources, capturing more accurately movements in the standard of living; third, it captures better non-monetary resources, particularly own-account transactions (self-supply) and transfers in kind from institutions and other households. The expenditure variable is constructed by the Peruvian National Institute of Statistics and takes into account the following consumption items received from any sources: food consumed inside and outside the household; clothing and shoes; housing rent; fuel, electricity and housing repairs; furniture and housing maintenance; health services and self-administered health care; transport and communications; leisure, amenities and education and cultural services; other goods and services. Excluded from the expenditure aggregate are public health or public education, the value of consumption services from durable goods and the consumption of water supplies taken from the river.

The consumption data are equivalised to accommodate differences in household size and structure, giving household expenditure per equivalent adult in 2005 purchasing power parity (PPP) terms. The analysis is carried out at the household, rather than individual, level and uses an equivalence scale constructed as $e = (A + \varphi C)^{\theta}$, where A and C are the numbers of adults and children in the household. We use parameters $\varphi = 0.65$ and $\theta = 0.70$, which are in line with Deaton and Zaidi's (2002) recommendations foran upper middle-income country like Peru. We use a time/location-secific poverty line consistent with the official poverty line provided by ENAHO to indicate non-extreme poverty. However, that poverty line is defined in relation to a per capita equivalence scale, so we adjust it by applying annual PPP factors and then multiplying by a factor defined as ratio of the regional means of the per capita and Deaton-Zaidi equivalence scales.

2.2 Patterns of panel response and nonresponse weights

Figure 1 shows the distribution of numbers of missing poverty observations within the 2007-11 observation window for the ENAHO panel. Non-response and attrition are a serious problem: 30% of the 2,031 households in the ENAHO panel have two or more missing observations and 18% have three or more missing. If we were to use a 5-year balanced panel, this would mean discarding 43% of panel members.

The ENAHO provides a combined weight variables designed to adjust for both unequal sampling rates in the survey design and non-response behaviour by panel members. The nonresponse component of this weighting variable is area-based. The country is divided by geographical region, county and degree of urbanisation into 178 areas, each of which is further partitioned into five socioeconomic strata. These units are then classified into five categories based on the 2005 population and housing census. The classification takes into account (*i*) dwelling conditions (based on floor and roof materials; (*ii*) household overcrowding; (*iii*) access to piped water and sewerage; and (*iv*) average childhood school attendance. Inverse probability response weights are then constructed sparately for each wave of the panel. The resulting weights are routinely used in research based on data from the ENAHO panel. Since our partial identification approach deals explicitly with non-response, we require only weights that adjust for unequal sampling rates, not for non-response. To construct these weights, we follow the procedure used by ENAHO, using the product of the inverse selection probabilities at each stage of the sample design. This takes into account the forecast population by age and sex for each month and district over the period of the panel (2007-2011).

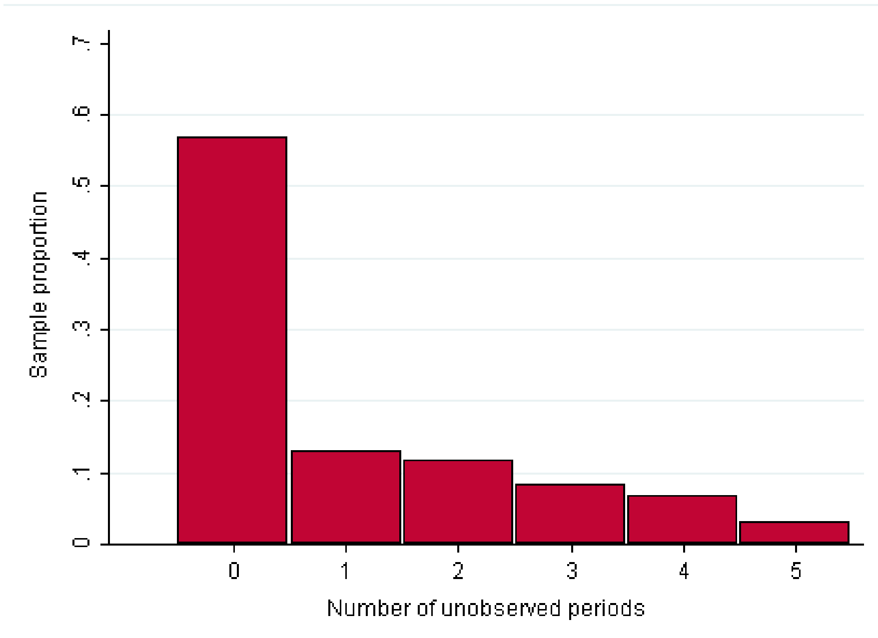

Figure 1 Distribution of numbers of missing observations in 5 waves
(ENAHO panel 2007-11; n = 2031 households, unweighted)

The well-established link between rural-urban migration and poverty reduction, together with the address-based design of the ENAHO panel, gives strong reason to question the MAR assumption which underlies conventional nonresponse weighting methods. There is indeed some evidence in the ENAHO data that the pattern of missing data might be related to the pattern of recurrent poverty. Define a crude poverty indicator for each household as the proportion of non-missing expenditure observations which are below the poverty line. Figure 2 shows the distribution of this indicator across households which are fully-observed and across households with some missing data. Under MAR, one would not expect any significant difference in the distribution of the proportion of time spent in poverty between the fully-observed and partially-observed cases, but the proportion of poverty-free households is far smaller for the fully-observed group. While this is not conclusive, it does underline concerns about the MAR assumption underpinning conventional analyses.

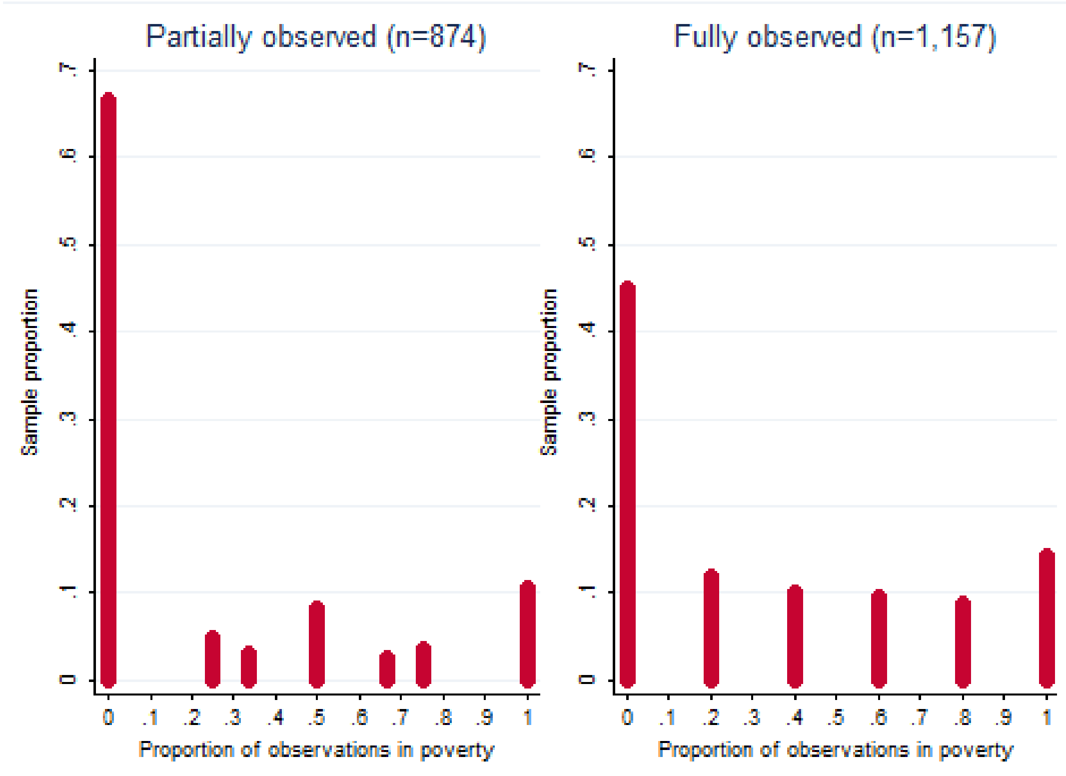

Figure 2 Distributions of the proportion of poverty observations in fully-observed and partially-observed households
(ENAHO panel 2007-11; $n = 2031$ households, design+non-response weights)

3 Measures of persistent poverty

Foster (2009) proposed a family of measures of chronic or persistent poverty. Here we consider the two members of that family which treat poverty as a binary state, together with a related measure used for monitoring purposes by the EU. We assume that panel data are available, covering a T-year observation period. Households are sampled randomly and an attempt is made to observe each household in periods $1...T$. If these attempts are successful, they yield a sample $\{Y_{it}; i = 1...n, t = 1...T\}$, where Y_t is a binary indicator of poverty in period t. In practice, some observations are missing and the pattern of 'missingness' may be non-ignorable, so that estimation of any poverty statistic from the balanced sub-panel of households observed in all T periods is potentially subject to bias. We begin by giving the simplest ("worst case") bounds, without using conditioning covariates or other external information. The simple nature of these measures makes it possible to work mainly with count variables summarising the poverty trajectory of the household within the observation window. All derivations are relegated to the appendix.

3.1 The persistent poverty headcount H

The simplest measure of persistent poverty is the headcount, Foster's (2009) H statistic, defined as the population proportion of households who experience poverty in at least C of the T periods covered by our panel: $H = Pr\left(\sum_{t=1}^{T} Y_t \geq C\right)$. This measure has a long history – for example Coe (1978) used a $C = T = 9$ headcount measure and Duncan et al (1984) used $C = 1, 5, 8, 10$ with $T = 10$ in their analyses of persistent poverty in the PSID panel. This type of measure fell out of favour after the influential work of Bane and Ellwood (1986), which focused instead on the initiation and duration of poverty spells. However, it remains widely used, for example in the official analysis of low-income dynamics in Britain (Department for Work and Pensions 2010). It is also frequently used in conjunction with the short panels that are typical in developing countries, where the Bane-Ellwood approach is harder to implement – see, for example, the review by Baluch and Hoddinott (2000) and recent work by Dercon and Porter (2011) and Dercon et al (2012) for Ethiopia.

Define P and N to be the numbers of observed periods of poverty and of non-poverty respectively for a generic household, and let P^*, N^* be the numbers of unobserved periods of poverty and non-poverty, where $P+N+P^*+N^* \equiv T$.[1] Write the joint probability distribution of the observed P, N as $f(P, N)$ and note that non-negativity and the inequalities $P+P^* \geq C$ and $P + P^* + N \leq T$ imply $0 \leq P \leq T$ and $0 \leq N \leq T - \max\{C, P\}$. Thus the probability of persistent poverty is:

$$H = Pr\left(P + P^* \geq C\right) = \sum_{P=0}^{T} \sum_{N=0}^{T-\max\{C,P\}} Pr\left(P^* \geq C - P | P, N\right) f\left(P, N\right) \qquad (1)$$

We show in the appendix that this probability must lie between the following sharp bounds:

$$L_H = Pr\left(P \geq C\right) \qquad (2)$$

$$U_H = Pr\left(N \leq T - C\right) \qquad (3)$$

[1] N^* is redundant henceforth, since it can be deduced with certainty from P, N and P^*, as T is a known constant. Note that we abstract from difficulties caused by exits from the population through death, emigration, etc.

The width of these bounds is $Pr(P < C, N \leq T - C)$, which is the proportion of ndeterminate' cases: those where observation is partial ($P + N < T$) and persistence cannot be ruled out ($N \leq T - C$). In practice, non-response and attrition rates rates are sufficiently high to make these bounds alarmingly wide.

3.2 The EU measure $\mathcal{E}$

A variant of the headcount has been adopted for policy monitoring by the EU. See European Commission 2009 for the full list of EU social exclusion indicators, Eurostat 2012 for official estimates and Jenkins and van Kerm (2011, 2012) for a critique. The EU measure gives a retrospective picture, taken from the viewpoint of the current period, defined as the most recent observation at time T. A household is defined to be in persistent poverty[2] if it is currently poor and was poor in at least two of the preceding three periods: in other words, it must be poor in at least three out of four successive periods, with one of those being the current period T. We generalise this by allowing the observation window and persistence threshold to be arbitrary. In formal terms, define D_T as an indicator of whether we observe the household in the current period T ($D_T = 1$) or not ($D_T = 0$). For the case to count as one of persistent poverty, we require there to be C periods of poverty in total and one of those to be period T, thus the measure is:

$$\mathcal{E} = Pr\left(P + P^* \geq C, Y_T = 1\right) \tag{4}$$

We show in the appendix that the sharp bounds are:

$$\begin{aligned} L_{\mathcal{E}} &= Pr\left(Y_T = 1, D_T = 1, P \geq C\right) && (5) \\ U_{\mathcal{E}} &= Pr\left(D_T = 0, N \leq T - C\right) + Pr\left(Y_T = 1, D_T = 1, P < C, N \leq T - C\right) \\ &+ Pr\left(D_T = 0, P \geq C\right) + Pr\left(Y_T = 1, D_T = 1, P \geq C\right) && (6) \end{aligned}$$

3.3 The duration-adjusted headcount K_0

A major drawback of the measures H and $\mathcal{E}$ is that they fail to distinguish between households that have different durations of poverty within the T-period observation window. Foster

[2] Or, in EU terminology *at risk of persistent poverty*

(2009) proposed the K_0 statistic, designed to improve on the crude headcount by incorporating the extent to which each persistently poor household exceeds the critical threshold.[3] It is defined as the headcount measure H multiplied by the mean number of periods of poverty conditional on poverty persistence as a proportion of T:

$$\begin{aligned} K_0 &= H \times E\left(P + P^* | P + P^* \geq C\right)/T \\ &= \frac{1}{T}\left[\sum_{P,N \in S} \sum_{P^*=\max\{0,C-P\}}^{T-P-N} \left[P + P^*\right] g\left(P^*|P,N\right) f\left(P,N\right)\right] \end{aligned} \tag{7}$$

where $g(\,.\,|\,P,N)$ is the conditional probability distribution of P^* and S is the set of integers P,N which allow the possibility of persistent poverty and thus satisfy $P,N \geq 0, P+N \leq T, N \leq T-C$. Partition S into three subsets S_1, S_2 and S_3:

Fully-observed, persistently poor:	$S_1 = \{P,N : P+N = T, P \geq C\}$
Part-observed, persistently poor:	$S_2 = \{P,N : P+N < T, P \geq C\}$
Part-observed, poverty status ambiguous:	$S_3 = \{P,N : N \leq T-C, P < C\}$

In this case, the sharp bounds are:

$$L_{K_0} = \frac{1}{T} \sum_{P,N \in (S_1 \cup S_2)} P f\left(P,N\right) \tag{8}$$

$$U_{K_0} = \frac{1}{T}\left[\sum_{P,N \in S_1} P f\left(P,N\right) + \sum_{P,N \in (S_2 \cup S_3)} (T-N) f\left(P,N\right)\right] \tag{9}$$

3.4 Evidence: worst-case bounds

We use simple Bayesian estimation, which has the advantage of coping easily with the more complex structure of the improved bounds developed later in section 4, without the technical difficulties of bias correction and construction of confidence sets (Tamer 2010). All bounds can be estimated from sample information in the form of a vector $\boldsymbol{f}$ of sample frequencies for (P,N), which has a multinomial distribution conditional on the vector of underlying population probabilities, $\boldsymbol{\pi}$, giving the likelihood:

$$l(\boldsymbol{f}\,|\,\boldsymbol{\pi}) \propto \prod_j \pi_j^{f_j} \tag{10}$$

[3] Bossert et al (2008) and Dutta et al (2011) propose differently-weighted versions of the headcount which take account of the continuity of poverty; Dickerson and Popli (2012) apply a combination of these measures to analyse the effect of poverty on child development. These modified measures could be bounded using an extension of our approach.

The natural conjugate prior for $\boldsymbol{\pi}$ is the Dirichlet distribution, which has the form $g(\boldsymbol{\pi}) \propto \prod_j \pi_j^{A\alpha_j}$, where α_j can be thought of as a prior estimate of π_j and A represents the amount of prior information, expressed in a form equivalent to sample size. The posterior distribution is also Dirichlet:

$$h(\boldsymbol{\pi} \mid \boldsymbol{f}) \propto \prod_j \pi_j^{f_j + A\alpha_j} \tag{11}$$

The mean and variance of the posterior for π_j are $\mu_j = (f_j + a\alpha_j)/(n + A)$ and $\mu_j[1 - \mu_j]/(n + A)$ respectively. We specify each α_j as $1/J$ where J is the number of elements in $\boldsymbol{\pi}$, and set $A = 50$, so that the prior contributes less than one eighth of the information used in estimation. This choice of α_j was made primarily to avoid zero probabilities in the posterior distribution corresponding to empty cells in the sample distribution, but the choice of α_j makes no discernible difference to the posterior distribution with A set at this level.[4]

Since the bounds are known functions of the true population probabilities $\boldsymbol{\pi}$, we can draw a large sample $\{L^{(r)}, U^{(r)}, s = 1...R\}$ from their posterior distribution and make probabilistic statements about their location. Table 1 presents posterior means and standard deviations for the worst-case bounds, together with conventional sample persistent poverty rates calculated from the subset of complete survey responses, using alternative weights designed to correct for sample design and for non-response and sample design jointly.

The bounds are very wide and essentially useless for the sort of policy monitoring role envisaged by the EU (Atkinson et al 2002), unless we are prepared to make additional strong assumptions about the pattern of change over time. Consider the example of the '3 years in 4' headcount, with posterior mean bounds $[0.114, 0.343]$. Suppose poverty declines dramatically, with the upper bound for a later 4-year period halved; this would give worst-case bounds for the *change* in persistent poverty of $[-0.229, 0.0575]$, which tell us very little about the direction and magnitude of change. The bounds are unhelpful from a statistical viewpoint too: although the non-response weights used in the ENAHO panel have a big positive effect on estimates of persistent poverty, the bounds tell us nothing about their validity since, in every case, the estimate weighted only for sample design and the estimate weighted for both design and non-response lie well within the identified interval.

[4] Rather than using raw frequencies for the f_j, we construct them as $f_j = n \sum_{i=1}^{n} w_i \xi_{ij}$, where w_i is the survey design weight and ξ_{ij} is a binary variable identifying cases in the jth cell of the (P, N) distribution.

Table 1 Estimated worst-case bounds
(posterior means and standard deviations)

Persistence threshold	Persistent poverty measure	Balanced panel sample $Pr(P > C \mid P + N = T)$ weighted for: design	design+response	Bayesian posterior means and standard deviations: Lower bound	Standard deviation	Upper bound	Standard deviation
4-wave panel (2007-10)							
$C = 2$	H	0.327	0.411	0.226	(0.009)	0.542	(0.011)
	$\mathcal{E}$	0.195	0.284	0.123	(0.007)	0.347	(0.010)
	K_0	0.237	0.320	0.156	(0.007)	0.399	(0.009)
$C = 3$	H	0.186	0.282	0.114	(0.007)	0.343	(0.010)
	$\mathcal{E}$	0.146	0.237	0.084	(0.006)	0.263	(0.009)
	K_0	0.166	0.255	0.100	(0.006)	0.300	(0.009)
5-wave panel (2007-11)							
$C = 2$	H	0.359	0.430	0.255	(0.009)	0.608	(0.011)
	$\mathcal{E}$	0.201	0.263	0.123	(0.007)	0.394	(0.011)
	K_0	0.237	0.313	0.156	(0.006)	0.412	(0.008)
$C = 3$	H	0.241	0.329	0.146	(0.008)	0.437	(0.011)
	$\mathcal{E}$	0.174	0.242	0.094	(0.006)	0.327	(0.010)
	K_0	0.190	0.273	0.112	(0.006)	0.343	(0.009)
$C = 4$	H	0.146	0.233	0.083	(0.006)	0.276	(0.010)
	$\mathcal{E}$	0.121	0.198	0.063	(0.005)	0.233	(0.009)
	K_0	0.133	0.215	0.074	(0.005)	0.247	(0.009)

4 Tighter bounds

These results give an overly pessimistic view of the prospects for clear inferences on poverty persistence because the worst-case identified sets include extreme regions that we know are highly implausible. A drawback of the partial identification approach is that, to make the bounds useful, we need to be able to exclude the complete set of implausibly extreme regions via simple constraints which rule them out without also excluding other more realistic regions. These constraints are often not easy to specify. Here we consider three sources of additional information that can be used to tighten the bounds.

4.1 External information on the per-period poverty rate

Worst-case bounds are wide because they allow some unkown probabilities of the form $Pr(P^* \geq C - P|P, N)$ to take any value from 0 to 1. But this unrestricted range contains some highly implausible regions. If, for example, we have observed no non-poverty and several periods of poverty for a particular household, then it is implausible to assume that the probability of crossing the persistent poverty threshold is zero. Conversely, if we observe a household with repeated non-poverty and no poverty, it is unreasonable to entertain a value close to 1 for the unknown probability $Pr(P^* \geq C - P|P, N)$. We pursue this idea using bounds developed by Zaigraev and Kaniovski (2010) under the assumption that the outcomes in the unobserved periods are a set of exchangeable Bernoulli trials with equal probabilities and unknown correlation.[5] Over the period 2007-11, official cross-section poverty rates in Peru fell fell from 0.383 to 0.258. To ensure that our prior information is conservative, we use 0.25 as the lower limit on the probability of poverty in any missing panel wave for a household with some observed poverty and no observed non-poverty, and 1-0.25 as the upper limit for a household with observed non-poverty and no observed poverty. Limits for other cases are set appropriately. The particular form of we use for these limits is specified in section A4 of the appendix.

Writing these context-specific a priori limits for the marginal probability of poverty in any of the unobserved periods, conditional on P, N as $[\epsilon_{PN}^{\min}, \epsilon_{PN}^{\max}]$, we show in the appendix that the modified bounds for the headcount measure are:

$$L_H^* = L_H + \sum_{P=0}^{C-1} \sum_{N=0}^{T-C} \max\left\{0, \frac{[T-P-N]\epsilon_{PN}^{\min} - C + P + 1}{T - N - C + 1}\right\} f(P, N) \tag{12}$$

$$U_H^* = U_H - \sum_{P=0}^{C-1} \sum_{N=0}^{T-C} \left[1 - \min\left\{1, \frac{[T-P-N]\epsilon_{PN}^{\max}}{C - P}\right\}\right] f(P, N) \tag{13}$$

where L_H and U_H are the original bounds defined by (2) and (3).

[5] Exchangeability is a strong assumption, so we have also computed estimates (available on request) where we confine the use of this additional information to cases with $P = C - 1, N = 0$ and $N = T - C - 1, P = 0$, where only a single additional observation of poverty or non-poverty respectively is required to classify the household unambiguously. In this special case, the Zaigraev-Kaniovski bounds are valid without the exchangeability assumption and the results are very similar to those presented below.

4.2 Instrumental variable restrictions

Assume we can observe a set of discrete characteristics $Z \in S_z$ believed a priori to be unrelated to persistent poverty in the sense that all of our measures $H, \mathcal{E}$ and K_0 are identical in all the subpopulations defined by points of support for Z. This means that the worst-case bounds $\{L(z), U(z)\}$ evaluated at any point of support z are valid for the overall persistent poverty measure, allowing us to narrow the interval (2)-(3) by using the maximal lower bound and minimal upper bound over S_z:[6]

$$\max_{z \in S_z} L(z) \leq H \leq \min_{z \in S_z} U(z) \tag{14}$$

If the distribution of Z is sufficiently coarse, it is possible to estimate $Pr(P \geq C|Z)$ and $Pr(N \leq T - C|Z)$ as sample proportions with the cells defined by Z, as we do in this application. Otherwise, the bounds could be estimated by fitting empirical models using estimation methods which are as flexible as possible. The obvious choice of instrumental variables Z is derived from information relating to survey fieldwork, such as interviewer characteristics (Nicoletti 2010, Nicoletti et al 2011). This requires assumptions that (i) the process of observing the household does not change its behaviour and poverty outcome and (ii) that poverty status has no influence on fieldwork procedures such as selection of interviewers. In our application, we use an instrument that identifies four groups: interviewer above/below median age × interviewer with/without supervisor job grade.

4.3 Monotone instrument restrictions

An alternative form of instrument, the monotone instrumental variable (MIV), was introduced by Manski (1995) and Manski and Pepper (2000, 2009). Here, we use a single discrete variable $W \in S_w$ with the chosen measure of persistent poverty evaluated at points of support w is known a priori to be weakly increasing in w. For any set of conditional bounds $\{L(w), U(w)\}$, the MIV bounds are:

$$E\left[\max_{w \leq W} L(w)\right] \leq H \leq E\left[\min_{w \geq W} U(w)\right] \tag{15}$$

[6] In the IV and MIV cases, we can also condition on other covariates X which are not used as instruments. The bounds (14) and (15) are then conditional on X; unconditional bounds are constructed by averaging with respect to the distribution of X. Introducing other covariates would tighten the bounds to some degree but, by increasing the dimensionality of the problem, they make it more difficult to avoid strong simplifying assumptions.

where the min and max are with respect to w and the expectations are with respect to the distribution of W. The variable W must be observable even for households that the survey never succeeds in interviewing, so this confines them in practice to characteristics of the locality and exterior of the dwelling.

5 Estimates of improved bounds

We compute the worst-case bounds and improved bounds for each of 50,000 replications drawn from the posterior distribution. The results are summarised in terms of the means and standard deviations of the posterior bounds (reported separately for the three measures in Tables 2-4) and in terms of the probability of coverage by the identification interval (in Figure 3). We construct improved bounds using the external information, IV and MIV approaches separately. We also construct a fourth set combining all three improvements, by taking the maximal lower bound and minimal upper bound at each replication.[7] We find a considerable difference in the character of the results between the headcount measure and the more elaborate EU and duration-adjusted measures. For the headcount, Table 2 shows results for various combinations of panel length T and persistence threshold C.

[7] It would be possible to combine the IV and MIV approaches by working from bounds conditioned on both Z and W, and combining the min and max operations in (14) and (15). We choose not to do this because it would greatly increase the number of sample cells containing very few observations, and raise concerns about robustness.

Table 2 Estimates of improved bounds: the headcount H
(posterior means and standard deviations)

Persistence threshold	Bounds	Balanced panel sample $Pr(P > C \mid P + N = T)$ weighted for: design	design+response	Bayesian posterior means and standard deviations: Lower bound	Standard deviation	Upper bound	Standard deviation
4-wave panel (2007-10)							
	Worst-case			0.226	(0.009)	0.542	(0.011)
	External			0.236	(0.009)	0.542	(0.011)
$C = 2$	IV	0.327	0.411	0.353	(0.035)	0.513	(0.015)
	MIV			0.265	(0.017)	0.503	(0.023)
	Combined			0.353	(0.035)	0.496	(0.019)
	Worst-case			0.114	(0.007)	0.343	(0.010)
	External			0.122	(0.007)	0.343	(0.010)
$C = 3$	IV	0.186	0.282	0.199	(0.028)	0.311	(0.014)
	MIV			0.142	(0.012)	0.325	(0.021)
	Combined			0.199	(0.028)	0.306	(0.014)
5-wave panel (2007-11)							
	Worst-case			0.255	(0.009)	0.608	(0.011)
	External			0.264	(0.009)	0.608	(0.011)
$C = 2$	IV	0.359	0.430	0.384	(0.033)	0.583	(0.015)
	MIV			0.308	(0.017)	0.571	(0.024)
	Combined			0.385	(0.033)	0.565	(0.020)
	Worst-case			0.146	(0.008)	0.437	(0.011)
	External			0.156	(0.008)	0.437	(0.011)
$C = 3$	IV	0.241	0.329	0.252	(0.029)	0.405	(0.015)
	MIV			0.179	(0.013)	0.416	(0.021)
	Combined			0.252	(0.029)	0.399	(0.015)
	Worst-case			0.083	(0.006)	0.276	(0.010)
	External			0.087	(0.006)	0.276	(0.010)
$C = 4$	IV	0.146	0.233	0.138	(0.022)	0.245	(0.013)
	MIV			0.105	(0.009)	0.264	(0.014)
	Combined			0.138	(0.021)	0.243	(0.012)

For the headcount, the IV approach based on interviewer characteristics is the most effective in narrowing the bounds, both by raising the lower bound considerably and lowering the upper bound to some extent. The MIV approach based on the number of local businesses also contributes significantly in some cases by lowering the upper bound. The use of external information contributes very little. For the headcount measure, the use of a weighted estimate calculated from the subset of households which are fully observed (the balanced panel) is quite successful when weights are used to adjust for both sample design and non-response. that weighted estimate always lies between the mean posterior lower and upper bound, whereas weighting for sample design tends to give an estimate below the mean posterior lower bound. Despite the use of all three types of additional information, the combined bounds remain sufficiently wide to make it virtually impossible to do the kind of policy monitoring over time that is envisaged by the EU (European Commission 2009).

The EU and duration-adjusted variants are more complex than the simple headcount, and they behave in rather different ways when additional information is introduced. The posterior mean bounds are presented in Tables 3 and 4 for the $\mathcal{E}$ and K_0 measures. External information on the marginal per-period poverty rate make a much bigger contribution here, lowering the upper bound considerably. The IV assumption is again the most effective in raising the lower bound. The conventional weighting adjustment for non-response does not work well. For the $\mathcal{E}$ measure (Table 3), the balanced panel estimate using weights that adjust for both sample design and non-response lies at or above the posterior mean combined upper bound for every combination of T and C. In contrast, the use of weights to adjust only for sample design gives balanced panel estimates lying within the mean indentified interval in most cases.

For the duration adjusted headcount K_0,the fully-weighted balanced panel estimate lies above the mean upper bound in a majority of the five T, C combinations, while the use of weights for sample design only generates an estimate below the mean lower bound in all but one case. Thus, the evidence for these more complex measures is that conventional weighting procedures cannot be relied upon, and that weights intended to remove non-response bias display a tendency to over-adjust. Of course, there is no guarantee that these findings for a particular panel would also apply in other measurement contexts.

Table 3 Estimates of improved bounds: the EU measure $\mathcal{E}$
(posterior means and standard deviations)

Persistence threshold	Bounds	Balanced panel sample $Pr(P > C \vert P + N = T)$ weighted for: design	design+response	Bayesian posterior means and standard deviations: Lower bound	Standard deviation	Upper bound	Standard deviation
4-wave panel (2007-10)							
	Worst-case			0.123	(0.007)	0.347	(0.010)
	External			0.127	(0.007)	0.248	(0.008)
$C = 2$	IV	0.195	0.284	0.201	(0.026)	0.308	(0.014)
	MIV			0.144	(0.011)	0.341	(0.016)
	Combined			0.201	(0.026)	0.248	(0.008)
	Worst-case			0.084	(0.006)	0.263	(0.010)
	External			0.088	(0.006)	0.199	(0.008)
$C = 3$	IV	0.146	0.237	0.126	(0.019)	0.232	(0.013)
	MIV			0.097	(0.009)	0.257	(0.013)
	Combined			0.126	(0.019)	0.199	(0.008)
5-wave panel (2007-11)							
	Worst-case			0.123	(0.007)	0.394	(0.011)
	External			0.125	(0.007)	0.263	(0.008)
$C = 2$	IV	0.201	0.263	0.196	(0.026)	0.365	(0.015)
	MIV			0.147	(0.009)	0.381	(0.018)
	Combined			0.197	(0.025)	0.263	(0.008)
	Worst-case			0.094	(0.006)	0.327	(0.010)
	External			0.098	(0.006)	0.222	(0.008)
$C = 3$	IV	0.174	0.242	0.148	(0.022)	0.294	(0.014)
	MIV			0.107	(0.008)	0.320	(0.016)
	Combined			0.149	(0.021)	0.222	(0.008)
	Worst-case			0.063	(0.005)	0.233	(0.009)
	External			0.064	(0.005)	0.170	(0.007)
$C = 4$	IV	0.121	0.198	0.100	(0.017)	0.199	(0.012)
	MIV			0.074	(0.007)	0.228	(0.013)
	Combined			0.100	(0.017)	0.170	(0.007)

Table 4 Estimates of improved bounds: the duration-adjusted headcount K_0 (posterior means and standard deviations)

Persistence threshold	Bounds	Balanced panel sample $Pr(P > C \| P + N = T)$ weighted for: design	design+response	Bayesian posterior means and standard deviations: Lower bound	Standard deviation	Upper bound	Standard deviation
4-wave panel (2007-10)							
	Worst-case			0.156	(0.007)	0.399	(0.009)
	External			0.172	(0.007)	0.330	(0.010)
$C = 2$	IV	0.237	0.320	0.247	(0.025)	0.373	(0.012)
	MIV			0.178	(0.010)	0.383	(0.017)
	Combined			0.247	(0.025)	0.330	(0.010)
	Worst-case			0.100	(0.006)	0.300	(0.009)
	External			0.107	(0.006)	0.230	(0.009)
$C = 3$	IV	0.166	0.255	0.172	(0.024)	0.272	(0.013)
	MIV			0.123	(0.011)	0.286	(0.016)
	Combined			0.173	(0.024)	0.230	(0.009)
5-wave panel (2007-11)							
	Worst-case			0.156	(0.006)	0.412	(0.011)
	External			0.177	(0.007)	0.348	(0.010)
$C = 2$	IV	0.237	0.313	0.245	(0.023)	0.385	(0.011)
	MIV			0.181	(0.009)	0.397	(0.016)
	Combined			0.245	(0.023)	0.347	(0.010)
	Worst-case			0.112	(0.006)	0.343	(0.009)
	External			0.124	(0.007)	0.249	(0.009)
$C = 4$	IV	0.190	0.273	0.193	(0.023)	0.314	(0.012)
	MIV			0.135	(0.009)	0.332	(0.016)
	Combined			0.193	(0.023)	0.249	(0.009)
	Worst-case			0.074	(0.005)	0.247	(0.009)
	External			0.079	(0.006)	0.178	(0.008)
$C = 4$	IV	0.133	0.215	0.126	(0.020)	0.218	(0.012)
	MIV			0.094	(0.008)	0.236	(0.013)
	Combined			0.126	(0.020)	0.178	(0.008)

5.1 'Credible' estimates

In Bayesian statistics, a credible set (at the $(1-\alpha)$ level) is a set of values for some unknown parameter which has $(1-\alpha)$ posterior probability. Extending the idea slightly, we define a credible set for any particular measure $M = H, \mathcal{E}$ or K_0 as the set $\{m : Pr([L_M, U_M] \ni m) \geq 1 - \alpha\}$, where the m are points in the unit interval, L_M and U_M are bounds relevant to the measure M and the probability is with respect to the posterior distribution of L_M, U_M. For example, a 90% credibility set for H is the set of values for the headcount which lie within the bounds at least 90% of the time when sampling from the posterior distribution. All the credibility sets are closed intervals in our case, and they can be read off from diagrams like Figure 3, which shows a plot of $Pr(m \in [L_M, U_M])$ for each of the three measures and for the five types of bound, in the case of a 4-wave panel and a persistence threshold of $C = 3$.

The combined use of external information and IV and MIV restrictions achieves a big reduction in the size of these credible regions. For example, in the case illustrated in Figure 3, the width of the identified interval falls from 0.207 to 0.052 for the headcount H; from 0.343 to 0.037 for the EU meaure $\mathcal{E}$; and from 0.180 to 0.012 for the duration-adjusted headcount K_0. These are remarkable improvements but, with the possible exception of K_0, they still leave the region of inherent uncertainty large enough to cause difficulties for the use of these measures for policy monitoring over time, given the size of changes likely to be experienced in practice.

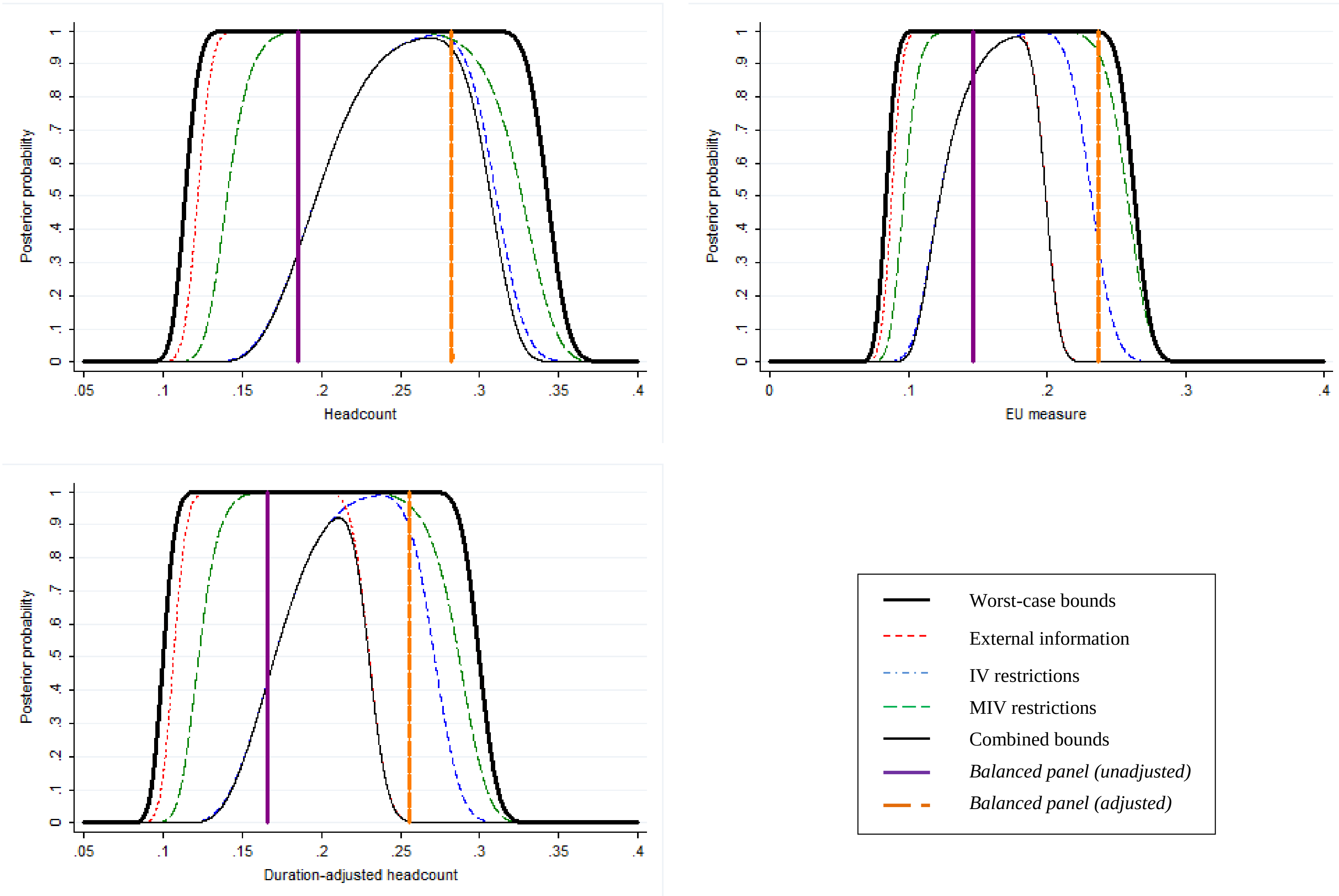

Figure 3 Posterior probabilities of coverage by bounds: $T = 4, C = 3$

6 Conclusions

There are three main conclusions. First, we have shown that the prevalence and pattern of missing data that is typical of household panel surveys produces much inherent uncertainty. If we impose no a priori restrictions and 'ask the data data to speak for themselves', the resulting bounds are very wide and completely useless for most analytical and policy purposes. This is true for all three of the persistent poverty measures we consider: the simple headcount, the EU risk of persistent poverty measure, and the duration-adjusted headcount.

Second, using the Peruvian ENAHO panel, we have shown that combined use of three types of plausible a priori information can be used to narrow the range of uncertainty considerably. They are: weak context-specific limits on the magnitude of the marginal single-period poverty risk; an IV-type assumption that the underlying poverty process is invariant to interviewer characteristics; and a monotonicity restriction relating poverty risk to an index of neighbourhood business enterprise. Of these types of additional information, the external limits are most effective in tightening the upper bound, and the IV assumption is the most effective in raising the lower bound.

Our third main finding is the unreliability of the common practice of using a balanced panel for multi-period poverty analysis, with use of conventional survey weights to remove non-response bias. In our application, neither the weights intended to compensate for sample design, nor the weights intended to adjust for sample design and non-response are able to produce estimates that fall within our improved bounds in most cases. This suggests that data are "missing not at random" – a situation that conventional survey weights are not designed to handle.

Although our final bounds are still rather too wide to allow reliable monitoring of poverty persistence as envisaged by the EU, our experience with the partial identification approach is sufficiently encouraging to suggest that there is scope for it to be used productively in a multi-period setting.

References

[1] Atkinson, A. B., Cantillon, B., Marlier, E. and Nolan, B. (2002). *Social Indicators. The EU and Social Inclusion*. Oxford: Oxford University Press.

[2] Baulch, R. and Hoddinott, J. (2000). Economic mobility and poverty dynamics in developing countries, *Journal of Development Studies* **36** 1-24.

[3] Bane, M. J. and Ellwood,D. T. (1986). Slipping into and out of poverty. The dynamics of spells. *Journal of Human Resources* **21**, 1-23.

[4] Bossert, W., Chakravarty, S. R. and DAmbrosio C. (2008). overty and Time, Working Paper No. 2008-87, ECINEQ, Society for the Study of Economic Inequality.

[5] Coe, R. D. (1978). Dependency and poverty in the short and long run, in G. J. Duncan and J. N. Morgan (eds.) *5000 American Families: Patterns of Economic Progress (vol. VI)*. Ann Arbor: Institute for Social Research, University of Michigan.

[6] Deaton, A. S. and Zaidi, S. (2002). *Guidelines for constructing consumption aggregates for welfare analysis*. Washington: World Bank Publications.

[7] Department for Work and Pensions (2010). *Low-Income Dynamics 1991-2008 (Great Britain)*. London: Department for Work and Pensions.

[8] Dercon, S., Hoddinott, J. and Woldehanna, T. (2012). Growth and chronic poverty: Evidence from rural communities in Ethiopia, *Journal of Development Studies* **48**, 238-253.

[9] Dercon, S. and Porter, C. (2011). A poor life? Chronic poverty and downward mobility in rural Ethiopia, 1994 to 2004. In R. Baulch (ed.), *Why Poverty persists: Poverty Dynamics in Asia and Africa*. Chronic Poverty Research Centre and UK department for International Development.

[10] Dercon, S. and Shapiro, J. S. (2007). Moving on, staying behind, getting lost: lessons on poverty mobility from longitudinal data. In D. Narayan and P. Petesch (eds.), *Moving Out of Poverty. Cross-Disciplinary Perspectives on Mobility*. Washington DC: World Bank.

[11] Dickerson, A. and Popli, G. (2012). Persistent poverty and children's cognitive development: Evidence from the UK Millennium Cohort Study. London: Institute for Education, Centre for Longitudinal Studies, Working paper 2012/2.

[12] Duncan, G. J., Coe, R. D. and Hill, M. S. (1984). The dynamics of poverty, in J. G. Duncan et al, *Years of Poverty, Years of Plenty*. Ann Arbot: Institute for Social Research, University of Michigan.

[13] Dutta, I., Roope, L. and Zank, H. (2011). On intertemporal poverty: affluence-dependent measures. University of Manchester: Economics Discussion Paper Series, EDP 1112.

[14] European Commission (2009). Portfolio of indicators for the monitoring of the European strategy for social protection and social inclusion – 2009 update. Brussels: European Commission (Employment, Social Affairs and Equal Opportunities Directorate) http://ec.europa.eu/social/BlobServlet?docId=3882langId=en (accessed 17 June 2013).

[15] Eurostat (2012). Persistent at-risk-of-poverty rates. Online database: http://epp.eurostat.ec.europa.eu/portal/page/portal/product_detailsdataset?p_product_code =TSDSC210 (accessed 23 September 2012).

[16] Foster, J. E. (2009). A class of chronic poverty measures. In T. Addison, D. Hulme and R. Kanbur (Eds.), *Poverty Dynamics: Interdisciplinary Perspectives*. Oxford: Oxford University Press.

[17] Jenkins, S. P. and van Kerm, P. (2011). Patterns of persistent poverty. University of Essex: ISER Working Paper no. 2011-30.

[18] Jenkins, S. P. and van Kerm, P. (2012). The relationship between EU indicators of persistent and current poverty. Bonn: IZA Discussion Paper no. 7071.

[19] Liao, Y. and Jiantg, W. (2010). Bayesian analysis in moment inequalty models, *Annals of Statistics* **38**, 275-316.

[20] Manski, C. F. (1995). *Identification Problems in the Social Sciences*. Cambridge MA: Harvard University Press.

[21] Manski, C. F. and Pepper, J. V. (2000). Monotone instrumental variables: with an application to the returns to schooling, *Econometrica* **68**, 997-1010.

[22] Manski, C. F. (2003). *Partial identification of probability distributions*, New York: Springer-Verlag.

[23] Moon, H. R. and Schorfheide, F. (2012). Bayesian and frequentist inference in partially identified models, *Econometrica* **80**, 755-782.

[24] Nicoletti, C. (2010). Poverty analysis with missing data: alternative estimators compared, *Empirical Economics* **38**, 1-22.

[25] Nicoletti, C., Peracchi, F. and Foliano, F. (2011). Estimating income poverty in the presence of missing data and measurement error, *Journal of Business Economics and Statistics* **29**, 61-72.

[26] Rosenzweig, M. R. (2003). Payoffs from panels in low-income countries: Economic development and economic mobility, *American Economic Review* **93**, 112-117. Thomas, D., Frankenberg, E. and Smith, J. P. (2001). Lost but not forgotten: attrition and follow-up in the Indonesian Family Life Survey, *Journal of Human Resources* **36**, 556-592.

[27] Zaigraev, A. and Kaniovski, S. (2010). Exact bounds on the probability of at least k successes in n exchangeable Bernoulli trials as a function of correlation coefficients, *Statistics and Probability Letters* **80**, 1079-1084.

Appendix: derivations

A1 The headcount measure

Split expression (1) into two components, using the fact that $Pr(P^* \geq C - P|P,N) = 1$ whenever $P \geq C$:

$$Pr(P + P^* \geq C) = \sum_{P=0}^{C-1} \sum_{N=0}^{T-C} Pr(P^* \geq C - P|P,N) f(P,N) + \sum_{P=C}^{T} \sum_{N=0}^{T-P} f(P,N)$$

The second component on the right-hand side can be identified from sample data, but the first component involves terms $Pr(P^* \geq C - P|P,N)$ which are unknown since P^* is unobserved. In the absence of further information about the distribution of $P^*|P,N$, lower and upper bounds (L_H and U_H) on the probability of persistent poverty are found by setting each term $Pr(P^* \geq C - P|P,N)$ to 0 and 1 respectively. Thus:

$$\begin{aligned}
L_H &= \sum_{P=C}^{T} \sum_{N=0}^{T-P} f(P,N) = Pr(P \geq C) \\
U_H &= \sum_{P=0}^{C-1} \sum_{N=0}^{T-C} f(P,N) + \sum_{P=C}^{T} \sum_{N=0}^{T-P} f(P,N) \\
&= Pr(P < C, N \leq T - C) + Pr(P \geq C) \\
&= Pr(P < C, N \leq T - C) + Pr(P \geq C, N \leq T - C) \\
&= Pr(N \leq T - C)
\end{aligned}$$

The penultimate step in the derivation of U_H follows because $P \geq C \Rightarrow N \leq T - C$.

A2 The EU measure

The EU measure can be written:

$$\begin{aligned}
\mathcal{E} &= \sum_{P=0}^{T} \sum_{N=0}^{T-\max\{C,P\}} \sum_{D_T=0}^{1} Pr(Y_T = 1 \mid D_T, P, N)\, Pr(P^* \geq C - P \mid P, N, Y_T = 1, D_T)\, Pr(D_T, P, N) \\
&= \sum_{P=0}^{C-1} \sum_{N=0}^{T-C} Pr(Y_T = 1 \mid D_T = 0, P, N)\, Pr(P^* \geq C - P \mid P, N, Y_T = 1, D_T = 0)\, Pr(D_T = 0, P, N) \\
&+ \sum_{P=0}^{C-1} \sum_{N=0}^{T-C} Pr(Y_T = 1 \mid D_T = 1, P, N)\, Pr(P^* \geq C - P \mid P, N, Y_T = 1, D_T = 1)\, Pr(D_T = 1, P, N) \\
&+ \sum_{P=C}^{T-1} \sum_{N=0}^{T-P} Pr(Y_T = 1 \mid D_T = 0, P, N)\, Pr(D_T = 0, P, N) \\
&+ \sum_{P=C}^{T} \sum_{N=0}^{T-P} Pr(Y_T = 1 \mid D_T = 1, P, N)\, Pr(D_T = 1, P, N)
\end{aligned}$$

All terms are potentially observable except $Pr(P^* \geq C - P \mid P, N, Y_T = 1, D_T = 0)$ and $Pr(Y_T = 1 \mid D_T = 0, P, N)$. Set these to 0 for the lower bound, and to 1 for the upper bound:

$$L_{\mathcal{E}} = \sum_{P=C}^{T} \sum_{N=0}^{T-P} Pr(Y_T = 1, D_T = 1, P, N)$$

$$U_{\mathcal{E}} = \sum_{P=0}^{C-1}\sum_{N=0}^{T-C}\left[Pr\left(D_T = 0, P, N\right) + Pr\left(Y_T = 1, D_T = 1, P, N\right)\right] + \sum_{P=C}^{T-1}\sum_{N=0}^{T-P}\left[Pr\left(D_T = 0, P, N\right) + Pr\left(Y_T = 1, D_T = 1, P, N\right)\right]$$

These define the bounds (5) and (6).

A3 The duration-adjusted headcount

Split expression (7) into sums over the sets $S_1 ... S_3$:

$$T \times K_0 = \sum_{P,N\epsilon S_1} Pf\left(P,N\right) + \sum_{P,N\epsilon S_2}\sum_{P^*=0}^{T-P-N}\left[P + P^*\right] g\left(P^*|P,N\right) f\left(P,N\right) + \sum_{P,N\epsilon S_3}\sum_{P^*=C-P}^{T-P-N}\left[P + P^*\right] g\left(P^*|P,N\right) f\left(P,N\right)$$

The only unknown terms are the conditional probabilities $g\left(P^*|P,N\right)$, and these can take any values consistent with the restrictions $g\left(P^*|P,N\right) \geq 0$ and $\sum_{P^*=0}^{T-P-N} g\left(P^*|P,N\right) = 1$. For the lower bound, choose the smallest possible value for the second and third terms by setting $g\left(P^*|P,N\right)$ equal to 1 for $P^* = 0$ and 0 for each $P^* > \max\{0, C-P\}$. For the upper bound, choose the largest possible values for the second and third terms by setting $g\left(P^* = T-P-N|P,N\right) = 1$ and $f\left(P^*|P,N\right) = 0$ for $P^* < T-P-N$, giving:

$$T \times L_{K_0} = \sum_{P,N\epsilon S_1} Pf\left(P,N\right) + \sum_{P,N\epsilon S_2} Pf\left(P,N\right)$$
$$T \times U_{K_0} = \sum_{P,N\epsilon S_1} Pf\left(P,N\right) + \sum_{P,N\epsilon S_2 \cup S_3} [T-N] f\left(P,N\right)$$

which are expressible as (8)-(9).

A4 External information

Zaigraev and Kaniovski (2010) (henceforth ZK) prove that, for a set of n exchangeable Bernoulli trials with equal probabilities p and unknown correlation, the probability of at least k occurrences satisfies the following bounds:

$$\lambda(n,k,p) \leq Pr(\text{at least } k \text{ occurrences}) \leq \upsilon(n,k,p)$$

where: $\lambda(n,k,p) = \max\{0, (np-k+1)/(n-k+1)\}$; $\upsilon(n,k,p) = \min\{1, np/k\}$; p is the household's marginal per-period poverty rate, assumed uniform over time; n is the number of unobserved periods; and k is the minimum number of periods of unobserved poverty required to reach the persistence threshold. Our a priori limits on p are $\epsilon_{PN}^{\min}$ and $\epsilon_{PN}^{\max}$ as set out in Table A1:

Table A1 Context-specific a priori limits on the period-specific marginal poverty rate

P	N	$\epsilon_{PN}^{\min}$	$\epsilon_{PN}^{\max}$
$[1...C-1]$	0	0.25	0.95
0	$[1...T-C]$	0.10	0.25
0	0	0.025	0.75
$P > N > 0$		0.20	0.75
$N \geq P > 0$		0.10	0.50

The headcount

To incorporate this further information, partition equation (1):

$$H = \sum_{P=0}^{C-1} \sum_{N=0}^{T-C} Pr\left(P^* \geq C - P | P, N\right) f\left(P, N\right) + \sum_{P=C}^{T} \sum_{N=0}^{T-P} f\left(P, N\right)$$

For the lower bound, instead of setting all unknown terms $Pr\left(P^* \geq C - P | P, N = 0\right)$ to zero, set them to the ZK lower bound, specifying $n = T - P - N, k = C - P$ and $p = \epsilon_{PN}^{\min}$, giving:

$$L_H^* = \sum_{P=0}^{C-1} \sum_{N=0}^{T-C} \lambda\left(T - P - N, C - P, \epsilon_{PN}^{\min}\right) f\left(P, N\right) + Pr\left(P \geq C\right)$$

For the upper bound, use the ZK upper limit on $Pr\left(P^* \geq C - P | P, N\right)$ rather than the extreme value of 1, giving:

$$U_H^* = \sum_{P=0}^{C-1} \sum_{N=0}^{T-C} \upsilon\left(T - P - N, C - P, \epsilon_{PN}^{\max}\right) f\left(P, N\right) + Pr\left(P \geq C\right)$$

These are improved bounds given in (12) and (13) above.

The EU measure

Write the EU measure as:

$$\begin{aligned}
\mathcal{E} \;=\; & \sum_{P=0}^{C-1} \sum_{N=0}^{T-C} Pr\left(Y_T = 1, P^* \geq C - P \mid P, N, D_T = 0\right) Pr\left(P, N, D_T = 0\right) \\
& + \sum_{P=0}^{C-1} \sum_{N=0}^{T-C} Pr\left(P^* \geq C - P \mid P, N, Y_T = 1, D_T = 1\right) Pr\left(Y_T = 1, D_T = 1, P, N\right) \\
& + \sum_{P=C}^{T} \sum_{N=0}^{T-P} Pr\left(Y_T = 1 \mid P, N, D_T = 0\right) Pr\left(P, N, D_T = 0\right) \;+\; Pr\left(Y_T = 1, D_T = 1, P \geq C\right)
\end{aligned}$$

The first term involves the unobservable probability $Pr\left(Y_T = 1, P^* \geq C - P \mid P, N, D_T = 0\right)$. Since $P^* = T - N - P$ implies $Y_T = 1$, this can be bounded below by $Pr(P^* = T - N - P | P, N, D_T = 0)$; an obvious upper bound is $Pr(P^* \geq C - P | P, N, D_T = 0)$, and both of these bounds can in turn be bounded using ZK. In the second term defining $\mathcal{E}$, the probability $Pr\left(P^* \geq C - P \mid P, N, Y_T = 1, D_T = 1\right)$ can be directly bounded using ZK. The third term involves $Pr\left(Y_T = 1 \mid P, N, D_T = 0\right)$, which is zero if $N = T - P$ and otherwise is the probability of a positive outcome in an exogenously chosen period and is thus bounded by $[\epsilon^{\min}_{PN}, \epsilon^{\max}_{PN}]$. Putting these together, gives the following bounds:

$$
\begin{aligned}
L^*_{\mathcal{E}} &= \sum_{P=0}^{C-1}\sum_{N=0}^{T-C} \lambda\left(T-N-P, T-N-P, \epsilon^{\min}_{PN}\right) Pr\left(P, N, D_T = 0\right) \\
&+ \sum_{P=0}^{C-1}\sum_{N=0}^{T-C} \lambda\left(T-N-P, C-P, \epsilon^{\min}_{PN}\right) Pr\left(Y_T = 1, D_T = 1, P, N\right) \\
&+ \sum_{P=C}^{T}\sum_{N=0}^{T-P} \epsilon^{\min}_{PN} Pr\left(P, N, D_T = 0\right) \quad + \quad Pr\left(Y_T = 1, D_T = 1, P \geq C\right) \\
U^*_{\mathcal{E}} &= \sum_{P=0}^{C-1}\sum_{N=0}^{T-C} \epsilon^{\max}_{PN} Pr\left(P, N, D_T = 0\right) \\
&+ \sum_{P=0}^{C-1}\sum_{N=0}^{T-C} \upsilon\left(T-N-P, C-P, \epsilon^{\max}_{PN}\right) Pr\left(Y_T = 1, D_T = 1, P, N\right) \\
&+ \sum_{P=C}^{T}\sum_{N=0}^{T-P} \epsilon^{\max}_{PN} Pr\left(P, N, D_T = 0\right) \quad + \quad Pr\left(Y_T = 1, D_T = 1, P \geq C\right)
\end{aligned}
$$

where, in the first term of the upper bound, we have used the fact that $\upsilon(T-P-N, T-P-N, \epsilon^{\max}_{PN}) \equiv \epsilon^{\max}_{PN}$.

The duration-adjusted headcount

Write the measure as:

$$
\begin{aligned}
T \times K_0 &= \sum_{P,N \in S_1 \cup S_2} P f\left(P, N\right) + \sum_{P,N \in S_2} \left[\sum_{P^*=0}^{T-P-N} P^* g\left(P^* | P, N\right)\right] f\left(P, N\right) \\
&+ \sum_{P,N \in S_3} P \left[\sum_{P^*=C-P}^{T-P-N} g\left(P^* | P, N\right)\right] f\left(P, N\right) + \sum_{P,N \in S_3} \left[\sum_{P^*=C-P}^{T-P-N} P^* g\left(P^* | P, N\right)\right] f\left(P, N\right)
\end{aligned}
$$

The unobserved elements are the three sums in square brackets. The first of these defines the conditional mean, $E(P^*|P, N)$, of the number of "successes" in a set of $T - P - N$ "trials", which is bounded above and below by $\left\{(T-P-N)\epsilon^{\min}_{PN}, (T-P-N)\epsilon^{\max}_{PN}\right\}$. The second defines the probability $Pr(P^* \geq C - P|P, N)$. The third is a partial sum which satisfies the following inequalities:

$$
(C-P)Pr(P^* \geq C-P|P,N) \leq \sum_{P^*=C-P}^{T-P-N} P^* g\left(P^*|P,N\right) \leq (T-P-N)Pr(P^* \geq C-P|P,N)
$$

The ZK bounds $\left\{\lambda\left(T-P-N, C-P, \epsilon^{\min}_{PN}\right), \upsilon\left(T-P-N, C-P, \epsilon^{\max}_{PN}\right)\right\}$ can be applied to $Pr(P^* \geq C - P|P, N)$, giving the following bounds for K_0:

$$
\begin{aligned}
L^*_{K_0} &= \frac{1}{T}\Bigg[\sum_{P,N \in S_1 \cup S_2} P f\left(P, N\right) + \sum_{P,N \in S_2} \epsilon^{\min}_{PN} P(T-P-N) f\left(P, N\right) \\
&+ \sum_{P,N \in S_2} P \lambda\left(T-P-N, C-P, \epsilon^{\min}_{PN}\right) f(P,N) + \sum_{P,N \in S_3} (C-P) \lambda\left(T-P-N, C-P, \epsilon^{\min}_{PN}\right) f(P,N)\Bigg] \\
U^*_{K_0} &= \frac{1}{T}\Bigg[\sum_{P,N \in S_1 \cup S_2} P f\left(P, N\right) + \sum_{P,N \in S_2} \epsilon^{\max}_{PN} P(T-P-N) f\left(P, N\right) \\
&+ \sum_{P,N \in S_2} P \upsilon\left(T-P-N, C-P, \epsilon^{\max}_{PN}\right) f(P,N) + \sum_{P,N \in S_3} (T-P-N) \upsilon\left(T-P-N, C-P, \epsilon^{\max}_{PN}\right) f(P,N)\Bigg]
\end{aligned}
$$