ECINEQ
Society for the Study of Economic Inequality

Working Paper Series

Robust "pro-poorest" poverty reduction with counting measures: the non-anonymous case

José V. Gallegos
Gastón Yalonetzky

ECINEQ WP 2014 - 351

ECINEQ 2014 - 351
December 2014

www.ecineq.org

Robust "pro-poorest" poverty reduction with counting measures: the non-anonymous case*

José V. Gallegos
Peruvian Ministry of Development and Social Inclusion, Peru

Gastón Yalonetzky
University of Leeds, UK

Abstract

When measuring poverty with counting measures, are there conditions ensuring that poverty reduction not only reduces the average poverty score further but also decreases deprivation inequality among the poor more, thereby emphasizing improvements among the poorest of the poor? In the case of a non-anonymous assessment, i.e. when we can track poverty experiences of the same individuals or households using panel datasets, we derive three conditions whose fulfillment allows us to conclude that multidimensional poverty reduction is more egalitarian in one experience vis-à-vis another one, for a broad family of poverty indices which are sensitive to deprivation inequality among the poor, and from an ex-ante conception of inequality of opportunity. We illustrate these methods with an application to multidimensional poverty in Peru before and after the 2008 world financial crisis

Keywords: Pro-poorest poverty reduction, multidimensional poverty, transition matrices.

JEL Classification: I32.

*The authors would like to thank Francesco Andreoli, Claudio Zoli, Olga Canto, Florent Bresson, Cesar Calvo, and participants at the 1st Conference of the Peruvian Economic Association, Lima, Peru, August 2014, and the 33rd IARIW General Conference, Rotterdam, Netherlands, August 2014, for their helpful comments. The usual disclaimer applies.

1 Introduction

While the “pro-poor" growth literature has traditionally worked with one continuous variable at a time, recently researchers have been interested in connecting the “pro-poor" growth concepts with non-monetary measures of well-being, and multidimensional poverty indices in particular. For instance, Berenger and Bresson (2012) provide dominance conditions to probe the “pro-poorness" of growth when well-being is measured jointly by continuous and discrete variables. Ben Haj Kacem (2013) measures the “pro-poorness" of growth in income when the initial conditioning situation is not income itself but a non-monetary multidimensional index of poverty or well-being. Boccanfuso et al. (2009) apply the now traditional “pro-poor" growth toolkit to assess changes in the individual scores of a non-monetary poverty composite index, where the weights are determined by multiple correspondence analysis (MCA). Since they use a vast number of indicators, their scores can take several values, thereby mimicking a continuous variable.

We pose a related question in the context of *multidimensional poverty counting measures*: What are the conditions under which a poverty reduction experience is more “pro-poorest" than another one *in a robust manner*? In other words, under which conditions does poverty reduction not only reduce the average poverty score further but also decrease deprivation inequality among the poor more, for a broad family of poverty indices that are sensitive to these distributional aspects? In a companion paper (Gallegos and Yalonetzky, 2014), we have answered this question for anonymous assessments in which we compare two cross-section datasets from different points in time, by adapting and extending methods developed by Lasso de la Vega (2010).

Meanwhile, when we have a panel dataset we can also perform a *non-anonymous* assessment of robust pro-poorest poverty reduction, in which we take into account the particular poverty transitions experienced by individuals or households.[1] For this purpose, in this paper we work with transition matrices and propose four second-order dominance conditions whose fulfillment allows us to conclude that multidimensional poverty reduction is more egalitarian in one poverty-transition experience vis-à-vis another one, for a broad family of poverty indices which are sensitive to deprivation inequality among the poor, and from an *ex-ante* conception of inequality of opportunity. [2]

[1]See Grimm (2007), for a thorough discussion of the distinction between anonymous and non-anonymous analysis of pro-poor growth in the continuous-variable context.

[2]See Fleurbaey and Peragine (2013) for an explanation of the distinction between *ex-ante* and *ex-post* inequality of opportunity.

The dominance rules focus on the distributions of expected deprivation scores conditioned on having certain values of the score in the initial period, i.e. inspired by the concept of “mobility as progressivity" put forward by Benabou and Ok (2001). The first, special rule works for the case in which both societies or time periods have the same initial *uniform* distributions of conditioning deprivation scores, i.e. the comparison is only based on differences in the vectors of conditional expected deprivation scores. The second, less restrictive rule works for the case in which both societies or time periods have the same initial (but not necessarily uniform) distributions of conditioning deprivation scores. The third, more general rule allows for differences (across time periods, or between societies) in the distributions of conditional expected deprivation scores, which in turn depend both on the initial distributions of conditioning scores, and on the vectors of possible values for the conditional expected scores. Finally, the four rule is an application of the third one, working with the *ergodic* probability distributions of expected conditional deprivation scores (i.e. the ergodic distribution from the transition matrix replaces the initial distribution of conditioning scores).

We illustrate the non-anonymous conditions using a yearly panel dataset from the Peruvian National Household Surveys spanning two periods: 2002-2006 and 2007-2010. In the former period, Peru experienced a commodity boom, which translated into high GDP growth rates, from 4 % in 2003 to 8.9 % in 2007, and a steady decrease in monetary poverty headcounts, from 58.7 % in 2004 to 42.4 % in 2007. However, between 2008 and 2013, Peru's economic performance was affected by the world economic situation: GDP growth fell from 9.8 % in 2008 to 0.9 % in 2009, and then stabilizing around 7 % between 2010 and 2012. Notwithstanding this fluctuation, monetary poverty levels kept decreasing steadily, from 37.3 to 27.8 %. But how did the Peruvian population fare in terms of non-monetary multidimensional poverty? We measure non-monetary poverty with wellbeing indicators corresponding to four dimensions: household education, dwelling material infrastructure, access to services, and vulnerability related to household dependency burden.

We search for robust rankings among the *poverty transitions* (2002-2004, 2004-2006, 2007-2008, and 2008-2010) according to their degree of *ex-ante* “pro-poorest" poverty reduction, i.e. the extent to which they reduce expected poverty while reducing inequality among the poor at the same time. When we implement the first condition, we find that the mobility matrix of 2002-2004 is the pro-poorest in the sense of yielding a set of expected conditional poverty experiences that second-order dominates the vector induced by all the other matrices. Then, the matrix of 2004-

2006 turns up as the second-best, since its expected distribution dominates those of 2007-2008 and 2008-2010, while being dominated by its predecessor's. Meanwhile the matrix of 2007-2008 turns up as the least desirable. In summary, the pre-crisis poverty transitions induced preferable distributions of expected poverty scores from a welfare-utilitarian point of view. When we implement the second condition, we find again that the matrix of 2002-2004 dominates all the others, except that of 2008-2010. By contrast, when we implement the third condition, the matrix of 2008-2010 dominates the two pre-crisis matrices, pointing to the importance of its more desirable initial distribution of conditioning deprivation scores. Finally, when we implement the fourth condition, based on ergodic distributions, we find that, again, the matrix of 2002-2004 prevails over the others, except for 2008-2010 against which it cannot be ordered. The latter matrix also dominates those of 2004-2006 and 2007-2008 in the case of the fourth condition.

The rest of the paper proceeds as follows: The next section presents our robust "pro-poorest" poverty-reduction conditions for non-anonymous assessments. First, it introduces the family of counting poverty measures for which the conditions are relevant and applicable, then it shows the three dominance conditions. The third section provides the empirical illustration on multidimensional poverty reduction in Peru. Finally, the paper concludes with some remarks.

2 Robust pro-poorest poverty reduction with counting measures

2.1 Inequality-sensitive poverty measures

Consider N individuals and D indicators of wellbeing. x_{nd} stands for the level of attainment by individual n on indicator d. If $x_{nd} < z_d$, where z_d is a deprivation line for indicator d, then we say that individual n is deprived in indicator d. In order to account for the breadth of deprivations, counting measures rely on individual deprivation scores which produce a weighted count of deprivations. If the weights are denoted by: $w_d \in [0,1] \subset \mathbb{R}_+ | \sum_{d=1}^{D} w_d = 1$, then the deprivation score for individual n is: $c_n \equiv \sum_{d=1}^{D} w_d \mathbb{I}(x_{nd} < z_d)$, where $\mathbb{I}$ is the indicator function. [3] Following Alkire and Foster (2011) we can also identify those multidimensionally poor with a flexible counting approach that compares each c_n against a multidimensional cut-off $k \in [0,1] \subset \mathbb{R}_+$, so

[3] Taking the value of 1 if the argument in parenthesis is true, otherwise it is equal to 0.

that person n is poor if and only if: $c_n \geq k$.

Our analysis focuses on a family of social poverty counting measures that are symmetric across individuals, additively decomposable (hence also subgroup consistent), scale invariant and population-replication invariant. If $p_n : c_n \times k \to [0,1] \in \mathbb{R}_+$ is the individual poverty measure, and $P : [0,1]^N \to [0,1]$ is the social poverty measure then our family is the following:

$$P = \frac{1}{N}\sum_{n=1}^{N} p_n \tag{1}$$

Our conditions of pro-poorest poverty reduction will also be useful for a broader family of subgroup consistent measures: $Q = H(P)$ as long as $H()$ is a strictly increasing, continuous function. For the sake of subgroup consistency, the weights must be set exogenously. Additionally we want P to fulfill the following key properties:

Axiom 1. *Focus (FOC): P should not be affected by changes in the deprivation score of a non-poor person as long as for this person it is always the case that: $c_n < k$.*

Axiom 2. *Monotonicity (MON): P should increase whenever c_n increases and n is poor.*

Axiom 3. *Progressive deprivation transfer (PROG): A rank-preserving transfer of a deprivation from a poorer individual to a less poor individual, such that both are deemed poor, should decrease P.*

In relation to the latter axiom, there are different approaches to capture sensitivity to deprivation inequality in the literature, although most of the approaches are virtually equivalent.[4] Axiom PROG is critical to the assessment of "pro-poorest" poverty reduction, as it forces social poverty indices to be sensitive to the distribution of deprivation across the poor, and to prioritize the wellbeing of the most jointly deprived among them.

In order to fulfill the above key properties, we narrow down the family of social poverty indices by rendering the functional form of p_n less implicit:

$$P = \frac{1}{N}\sum_{n=1}^{N} \mathbb{I}(c_n \geq k) g(c_n), \tag{2}$$

where $\mathbb{I}(c_n \geq k)$ is the Alkire-Foster poverty identification function that also secures the fulfillment of FOC; and $g : c_n \to [0,1]$, such that: $g(0) = 0$, $g(1) = 1$, $g' > 0$ and $g'' > 0$. The function g captures the intensity of poverty, which is understood as

[4]For a comparative review of these approaches see Silber and Yalonetzky (2013). A different framework is provided by Alkire and Seth (2014).

number of deprivations in the counting approach. Several examples of g have been proposed by Chakravarty and D'Ambrosio (2006).

2.2 The non-anonymous case based on the distributions of expected deprivation scores

In the counting approach, there is only one vector of possible values of c_n for each particular choice of deprivation lines and weights. We define that vector: $V := (v_1, v_2, ..., v_l)$, and note that $v_1 = 0$ (the case of someone without any deprivation), and $v_l = 1$ (the case of someone deprived in every indicator). Moreover it is easy to show that the maximum number of possible values is given by: $\max l = \sum_{i=0}^{D} \binom{D}{i}$. In the particular, but common, case of equal weights ($w_d = \frac{1}{D}$), the number of possible values is much smaller: $l = D + 1$. Hence the distribution of c_n in the sample is bound to be discrete, as there will be several individuals for every value of c_n.

In the non-anonymous case we can track the experience of each individual across periods with a panel dataset. More precisely we can construct a transition matrix with the social probabilities of attaining a particular deprivation score in the final year of the period, conditional on having had a specific deprivation score in the initial year. Then we can compute the expected value of the deprivation score conditional on a given value of the deprivation score in the initial year, by adding the products of the conditional probabilities of attaining each score in the final year times the score itself.

Thus we have as many conditional expected values of the score as the values for the score, i.e. l. Then we can provide social evaluations of the distribution of the conditional expected values of the scores. For instance, we may require properties similar to MON and PROG in the social evaluation of expected values. Hence, inspired by Benabou and Ok (2001), we can implement an ex-ante non-anonymous assessment of robust "pro-poorest" poverty reduction. In this assessment, we compare different transition matrices and rank them in terms of their capacity to reduce poverty, prioritizing reductions in the expected deprivation score of those who start the poorest.

If this assessment is applied to samples of parents and their adult offspring (so that the initial period corresponds to the former, and the final period to the latter), or at least to relatively long periods (e.g. several years), then it can also become an analysis of ex-ante inequality of opportunity (i.e. as long as we normatively posit that poverty prospects should not depend on past poverty experiences over which there is little or no control).

2 ROBUST PRO-POOREST POVERTY REDUCTION WITH COUNTING MEASURES

Let c_n^t be the score of individual n in period t. The probability of attaining a particular score in period t conditional on a specific score attained in period $t-1$ is defined as: $m_{i|j} = \Pr[c_n^t = i | c_n^{t-1} = j]$. The array of all these probabilities (i.e. from $m_{0|0}$ to $m_{1|1}$) constitutes a transition matrix M. If the number of values for the deprivation score (given a choice of weights and deprivation lines) is l then the transition matrix is an l-dimension square matrix. For any initial score value in period $t-1$ the conditional expected score in period t is:

$$E[c_n^t | v_j] = 0 \times m_{0|v_j} + v_2 \times m_{v_2|v_j} + v_3 \times m_{v_3|v_j} + \ldots + 1 \times m_{1|v_j}, \tag{3}$$

where the sum in 3 has l elements. Note also that, unlike c_n, $E[c_n^t|v_j]$ for any j can, in principle, take any value in the real interval $[0,1]$, even when weights and deprivation lines are fixed, because it also depends on the elements of M. The vector of expected deprivation scores is: $E := (E[c_n^t|0], E[c_n^t|v_2], \ldots, E[c_n^t|1])$. Consider also an l-dimensional vector Π containing the probability distribution of scores in period $t-1$: $\Pi := [\pi(0), \pi(v_2), \ldots, \pi(1)]$. $\pi(v_i) \equiv \Pr[c_n^{t-1} = v_i]$ and $\sum_{i=1}^{l} \pi(v_i) = 1$.

Now for the sake of presenting the conditions below, it is useful to rely on the following assumption:

Assumption 1. $E[c_n^t|1] \geq E[c_n^t|v_{l-1}] \geq \ldots \geq E[c_n^t|v_2] \geq E[c_n^t|0]$.

Assumption 1 states that the expected deprivation score does not decrease whenever the conditioning score value in the initial period increases. This assumption is not really necessary to derive the "pro-poorest" conditions below, but it substantially simplifies their presentation. Moreover, the assumption is likely to manifest regularly in empirical applications. In fact, all the transition matrices in our illustration satisfy assumption 1.[5]

For the conditions below, we also need to use a reverse generalized Lorenz (RGL) curve of the distribution of expected deprivation scores at the beginning and at the end of the time period. We define firstly the RGL curve for the distribution of expected deprivation scores not weighted by their relative frequency in the population. Hence an RGL curve is a function $L : [0,1] \to [0,1]$ that maps from a cumulative number s of the highest values in the vector of conditional expected deprivation scores, i.e. *ranked*

[5]Note that assumption 1 is much less stringent than imposing monotonicity on the matrices. A monotone transition matrix is characterized by consecutive first-order stochastic dominance between its adjacent columns (or rows). For example, in the case of M, monotonicity would require: $\sum_{i=1}^{k} m_{v_i|v_j} \geq \sum_{i=1}^{k} m_{v_i|v_{j+1}}$ $\forall j \in [1,2,\ldots,l-1], k \in [1,2,\ldots,l]$. Clearly, a monotone transition matrix suffices to uphold assumption 1, but it is not necessary.

from the highest to the lowest values of $E[c_n^t|v_j]$, to the incomplete average of $E[c_n^t|v_j]$, i.e. the sum of all the highest s scores, $E[c_n^t|v_j]$, divided by l:

$$L(s) = \frac{1}{l}\sum_{j=1}^{s} E[c_n^t|v_{l-j+1}] \quad s = 1, 2, ..., l. \tag{4}$$

Now using superscripts to denote populations where appropriate, we propose the following theorem:

Theorem 1. $\frac{1}{l}\sum_{j=1}^{l} g(E^A[c_n^t|v_j]) < \frac{1}{l}\sum_{j=1}^{l} g(E^B[c_n^t|v_j])$ *for all convex, strictly increasing, continuous functions* g*, if and only if* $L^A(s) \leq L^B(s) \quad \forall s \in [1, 2, ..., l] \quad \wedge \exists s | L^A(s) < L^B(s)$.

Proof. This theorem is similar in spirit to theorem 2 by Shorrocks (1983); the difference being that the former uses a RGL curve, whereas the latter involves a generalized Lorenz curve. However since Muirhead's theorem (e.g. see Marshall et al., 2011, p. 7-8) applies equally to values (of incomes, expected deprivation scores, etc.) ranked in descending or in ascending order, both results can be obtained. A full-fledged proof is available on request.[6] □

When Theorem 1 holds, M^A induces a stronger reduction in poverty than M^B, in terms of prioritizing the expected deprivation scores of those who start with higher scores in $t-1$ (under assumption 1).

2.2.1 From the union approach to other poverty identification approaches

Note that Theorem 1 assumes a union approach to poverty identification (i.e. k lower than the minimum deprivation weight in the poverty identification function $\mathbb{I}(c_n \geq k)$). However, the theorem can be restricted and applied to less lenient poverty identification criteria. The route to follow is to censor all scores whose value is below a chosen k_{min}, i.e. replacing c_n with c_n^* such that: $c_n^* = 0$ if $c_n < k_{min}$; otherwise $c_n^* = c_n$. Then the rest of the analysis proceeds as established above, noting that some cells in the transition matrices will be merged.

[6] RGL curves have also been used by Lasso de la Vega (2010) and Chakravarty and Zoli (2009) in the case of integer variables. In their applications they compare the distributions of deprivation scores, c_n, between A and B, but both sharing, effectively, the same vectors V. Therefore any social welfare differences between A and B are driven by differences in Π. By contrast, theorem 1 works with E^A and E^B, the vectors of expected deprivation scores, which are generally not identical, while implicitly assuming $\pi^A(i) = \pi^B(i) = \frac{1}{l} \ \forall i \in [1, 2, .., l]$. Theorem 2 below works with E and, more generally, $\pi^A(i) = \pi^B(i) \forall i \in [1, 2, .., l]$. Again, in this case, the vectors of expected deprivation scores are not identical, and in fact they drive any differences in the RGL curves, since otherwise the initial distributions are identical. Finally, theorem 3 below works with E and Π, and in both cases, A and B can be different.

2.3 The non-anonymous case with equal initial distributions of deprivation scores

Theorem 1 focuses exclusively on the values that the expected deprivation scores can take in each population, thereby neglecting the proportions of people who face each possible expected deprivation score. If we want the welfare evaluation of expected deprivation scores to be sensitive to their distribution, then we redefine the social evaluation function for the expected deprivation scores so that now we have: $\sum_{j=1}^{l} \pi(v_j) g(E[c_n^t|v_j])$. One starting point is to consider $\pi^A(i) = \pi^B(i) \forall i \in [1, 2, .., l]$, generally (i.e. now $\pi^A(i) = \pi^B(i) = \frac{1}{l} \forall i \in [1, 2, .., l]$ is just one particular case among many others). In this case of equal and general initial distributions of deprivation scores, it is easy to show that Theorem 1 does not work anymore. Instead we require the following Theorem 2:[7]

Theorem 2. $\frac{1}{l}\sum_{j=1}^{l} \pi(v_j) g(E^A[c_n^t|v_j]) < \frac{1}{l}\sum_{j=1}^{l} \pi(v_j) g(E^B[c_n^t|v_j])$ *for all convex, strictly increasing, continuous functions* g*, and for every possible* Π*, if and only if* $E^A[c_n^t|v_j] < E^B[c_n^t|v_j] \quad \forall j \in [1, 2, ..., l]$.

Proof. The sufficiency of $E^A[c_n^t|v_j] < E^B[c_n^t|v_j] \quad \forall j \in [1, l]$ is easy to ascertain, given the strict monotonicity properties of g. For the necessity, take each and every initial probability to its limit of 1, and it becomes clear that vector dominance of E^A over E^B is required to ensure that $\frac{1}{l}\sum_{j=1}^{l} \pi(v_j) g(E^A[c_n^t|v_j]) < \frac{1}{l}\sum_{j=1}^{l} \pi(v_j) g(E^B[c_n^t|v_j])$ for every possible Π. □

Interestingly, Theorem 2 states that M^A induces a stronger reduction in poverty than M^B (in the terms mentioned above), whenever the distributions of initial deprivation scores are identical, if and only if *for every initial deprivation score*, the respective conditional expected deprivation score in A is lower than in B; i.e. only provided there is vector dominance of E^A over E^B.

2.4 The non-anonymous case with different initial distributions of deprivation scores

We could also justify taking into account the actual proportions of people who would be facing each possible expected deprivation score, in each period (or society), i.e. the actual Π. For instance it could be the case that two societies have the same vectors of conditional expected deprivation scores, except that the highest expected score in

[7] I would like to thank Francesco Andreoli for helping me derive this result.

society A is higher than in B. Both theorems 1 and 2 would judge the distribution of B to be preferable. However, if now the distribution Π^A is different from Π^B, and in fact A has a smaller proportion of its population facing the highest expected score, then a priori we could not be sure anymore that any social welfare function based on any individual evaluation function g (as in Theorem 1 or 2) would rank B above A.

In this subsection we derive an alternative to Theorem 2 relevant to social poverty comparisons in which Π^A and Π^B are different. We consider the following reverse generalized Lorenz curve:

$$L_\pi(s) = \sum_{j=1}^{s} E[c_n^t|v_{l-j+1}]\pi(v_{l-j+1}), \quad s = 1, 2, ..., l. \tag{5}$$

Then theorem 3 states the following:

Theorem 3. $\sum_{j=1}^{l} \pi^A(v_j)g(E^A[c_n^t|v_j]) < \sum_{j=1}^{l} \pi^B(v_j)g(E^B[c_n^t|v_j])$ *for all convex, strictly increasing, continuous functions* g, *if and only if* $L_\pi^A(s) \leq L_\pi^B(s) \quad \forall s \in [1, 2, ..., l] \quad \wedge \exists s|L_\pi^A(s) < L_\pi^B(s)$.

Proof. Similar to the proof used for Theorem 1, but now the distributions Π^A and Π^B need to be considered. Available upon request. □

When Theorem 3 holds, M^A, *together with* Π^A, yields a distribution of expected conditional deprivation scores characterized by lower average expected deprivation scores and less inequality (in the sense captured by PROG) vis-a-vis the distribution of expected scores generated by M^B in combination with Π^B.

There is also an interesting sufficient condition guaranteeing that $L_\pi^A(s)$ is never above $L_\pi^B(s)$, and is strictly below for at least one value of s. In order to introduce it we need to define the poverty headcount *in the initial period*:

$$H(k) \equiv \frac{1}{N} \sum_{n=1}^{N} \mathbb{I}(c_n \geq k) \tag{6}$$

Now we get the following proposition:

Proposition 1. *If* $(E^A[c_n^t|v_{l-j+1}] - E^A[c_n^t|v_{l-j}]) \leq (E^B[c_n^t|v_{l-j+1}] - E^B[c_n^t|v_{l-j}]) \forall j \in [1, 2, ..., l-1] \wedge E^A[c_n^t|v_1] \leq E^B[c_n^t|v_1]$ *(with at least one of the former inequalities being strict),* and $H^A(v_i) \leq H^B(v_i) \ \forall i \in [1, 2, ..., l] \ \wedge \exists i|H^A(v_i) < H^B(v_i)$, *then:* $L_\pi^A(s) \leq L_\pi^B(s) \quad \forall s \in [1, 2, ..., l] \quad \wedge \exists s|L_\pi^A(s) < L_\pi^B(s)$.

Proof. Firstly, using summation by parts (Abel's formula) we can show that: $L_\pi^A(s) = \sum_{j=1}^{s-1} H(v_{l-j+1})[E^A[c_n^t|v_{l-j+1}] - E^A[c_n^t|v_{l-j}]] + H(v_{l-s+1})E^A[c_n^t|v_{l-s+1}]$. Secondly, we can also show that if $(E^A[c_n^t|v_{l-j+1}] - E^A[c_n^t|v_{l-j}]) \leq (E^B[c_n^t|v_{l-j+1}] - E^B[c_n^t|v_{l-j}]) \forall j \in [1,2,...,l-1] \wedge E^A[c_n^t|v_1] \leq E^B[c_n^t|v_1]$, then it must be the case that: $E^A[c_n^t|v_j] \leq E^B[c_n^t|v_j]\ \forall j \in [1,2,...,l]$. Finally, the rest follows by inspection. □

Proposition 1 states that if the absolute gaps between expected scores in A are never larger than in B (and they are strictly lower for at least one gap), *and* the poverty headcounts in period $t-1$ are never higher in A than in B (and they are strictly lower for at least one score value) then any social poverty function described in Theorem 3 ranks A favourably with respect to B.

2.4.1 The non-anonymous case with different ergodic distributions of deprivation scores

Some mobility assessments based on transition matrices often resort to the ergodic, or equilibrium distributions (e.g. Kremer et al., 2001). In the non-anonymous case of robust 'pro-poorest' poverty reduction we can also apply directly Theorem 3 to ergodic distributions, i.e. replacing Π with $\hat{\Pi}$, where $\hat{\Pi} := (\hat{\pi}(0), \hat{\pi}(v_2), ..., \hat{\pi}(1))$ is the ergodic distribution of deprivation scores. The idea would be that, after each transition, the spot initial distribution Π changes, until equilibrium is reached (provided that the transition matrix fulfills some regularity conditions). So the ergodic distribution help us conduct robust poverty comparisons in expected deprivation scores when the transition matrix does not change its distribution anymore.

Interestingly, the following proposition provides a sufficient condition, based only on transition probabilities, whose fulfillment guarantees that M^A yields an *ergodic* distribution of expected conditional deprivation scores characterized by lower average expected deprivation scores and less inequality (in the sense captured by PROG) vis-a-vis the ergodic distribution of expected scores generated by M^B.

Proposition 2. $\sum_{j=1}^{l} \hat{\pi}^A(v_j) g(E^A[c_n^t|v_j]) < \sum_{j=1}^{l} \hat{\pi}^B(v_j) g(E^B[c_n^t|v_j])$ *for all convex, strictly increasing, continuous functions g, if* $\forall j \in [1,2,...,l] : \sum_{i=1}^{q} m^A_{v_i|v_j} \geq \sum_{i=1}^{q} m^B_{v_i|v_j}\ \forall q \in [1,2,...,l] \wedge \exists q | \sum_{i=1}^{q} m^A_{v_i|v_j} > \sum_{i=1}^{q} m^B_{v_i|v_j}$.

Proof. 1. Adapting a result by Dardanoni (1995), we can show that: $\forall j \in [1,2,...,l] : \sum_{i=1}^{q} m^A_{v_i|v_j} \geq \sum_{i=1}^{q} m^B_{v_i|v_j}\ \forall q \in [1,2,...,l]\ \wedge \exists q | \sum_{i=1}^{q} m^A_{v_i|v_j} > \sum_{i=1}^{q} m^B_{v_i|v_j}$ implies $\hat{H}^A(v_s) < \hat{H}^B(v_s)\ \forall s \in [2,...,l]$ (by definition: $\hat{H}^A(v_1) = \hat{H}^B(v_1) = 1$).

2. Let $F_{v_q|v_j} \equiv \sum_{i=1}^{q} m_{v_i|v_j}$ for $1 \leq q \leq l$ be a cumulative transition probability. Now if we sum 3 by parts we get: $E[c_n^t|v_j] = -\sum_{i=1}^{l-1} F_{v_i|v_j}[v_{i+1} - v_i] + 1$. Hence if $\forall j \in [1,2,...,l] : \sum_{i=1}^{q} m^A_{v_i|v_j} \geq \sum_{i=1}^{q} m^B_{v_i|v_j}\ \forall q \in [1,2,...,l]\ \wedge \exists q | \sum_{i=1}^{q} m^A_{v_i|v_j} > \sum_{i=1}^{q} m^B_{v_i|v_j}$, then it must be the case that: $E^A[c_n^t|v_j] < E^B[c_n^t|v_j]\ \forall j \in [1,2,...,l]$.

3. Now let $\mathbb{A} \equiv \sum_{j=1}^{l} \hat{\pi}^A(v_j) g(E^A[c_n^t|v_j])$. We can express it the following way: $\mathbb{A} = \sum_{j=1}^{l} \hat{\pi}^A(v_j)[g(E^B[c_n^t|v_j]) + (g(E^A[c_n^t|v_j]) - g(E^B[c_n^t|v_j]))]$. Then the social welfare difference between A and B is: $\mathbb{A} - \mathbb{B} = \sum_{j=1}^{l}[\hat{\pi}^A(v_j) - \hat{\pi}^B(v_j)] g(E^B[c_n^t|v_j]) + \sum_{j=1}^{l} \hat{\pi}^A(v_j)[g(E^A[c_n^t|v_j]) - g(E^B[c_n^t|v_j])]$.

4. Integrating by parts the first element on the right-hand side of the expression for $\mathbb{A} - \mathbb{B}$ in point 3, we now get: $\mathbb{A} - \mathbb{B} = \sum_{j=2}^{l}[\hat{H}^A(v_j) - \hat{H}^B(v_j)][g(E^B[c_n^t|v_j]) - g(E^B[c_n^t|v_{j-1}])] + \sum_{j=1}^{l} \hat{\pi}^A(v_j)[g(E^A[c_n^t|v_j]) - g(E^B[c_n^t|v_j])]$

5. The rest follows by inspection, bringing together the results in points 1 and 2, plus the properties of g and assumption 1.

□

Even though its fulfillment guarantees a robust comparison of ergodic distributions of expected deprivation scores, one advantage of Proposition 2 is that, it does not require the actual computation of ergodic distributions, as it relies solely on transition probabilities. Hence the proposition can actually be tested with standard procedures, as proposed in the next section.

3 Statistical inference

3.1 Test of Theorem 2

In the empirical application, we test the condition from Theorem 2 after that for Theorem 1. However, it is more convenient to begin this section with an explanation of the test for Theorem 2. Different tests are possible, but we implement one in which a null hypothesis of $E^A[c_n^t|v_j] \leq E^B[c_n^t|v_j] \ \forall j \in [1, 2, ..., l]$ is set against an alternative whereby $\exists j | E^A[c_n^t|v_j] > E^B[c_n^t|v_j]$. If we do not reject the null we can state that either A vector-dominates B or the two vectors perfectly overlap, which implies that expected poverty in A can never be above B for the given choice of weights and deprivation lines, and for any poverty function considered in Theorem 2. Alternatively, rejecting in favour of the alternative means that "A does not dominate B", i.e. either A is dominated by B or the two distributions of expected deprivation scores cross (which in turn implies that the poverty comparison is sensitive to the choice of poverty functions, even for a given set of weights and deprivation lines).

In practice, we have a joint intersection null hypothesis: $Ho: E^A[c_n^t|v_j] = E^B[c_n^t|v_j] \ \forall j \in [1, 2, ..., l]$ against a union alternative $Ha: \exists j | E^A[c_n^t|v_j] > E^B[c_n^t|v_j]$. For that purpose, and considering that A and B are independently distributed, we construct the following statistics:

$$T(j) = \frac{E^A[c_n^t|v_j] - E^B[c_n^t|v_j]}{\sqrt{\frac{\sigma^2_{E^A}(j)}{\pi^A(v_j)N^A} + \frac{\sigma^2_{E^B}(j)}{\pi^B(v_j)N^B}}}, \tag{7}$$

where:

$$\sigma^2_{E^A}(j) \equiv V\Omega^A(v_j)V', \tag{8}$$

and $\Omega^A(v_j)$ is an $l \times l$ covariance matrix such that the diagonal element i is: $\Omega^A_{ii}(v_j) \equiv m^A_{i|j}(1 - m^A_{i|j})$ and the off-diagonal element for row i and column k consists of: $\Omega^A_{ik}(v_j) \equiv -m^A_{i|j}m^A_{k|j}$. (See e.g. Formby et al., 2004)

Then we test $Ho : T(j) = 0$ against $Ha : T(j) > 0$ for every value of j. Given the requirements of Theorem 2, we conclude that A does not dominates B if there is at least one j for which $T(j) > T_\alpha$, where T_α is the right-tail critical value for a one-tailed "z-test" corresponding to a level of significance α. Since we test multiple comparisons, the actual size of the whole test is not α. Under reasonable assumptions, it can be shown that it is $\beta = \sum_{i=1}^{l}[l-i+1]\alpha^i(-1)^{i-1}$. With $l = 5$, we choose $\alpha = 0.01$, so that $\beta \approx 0.05$.

3.2 Test of Theorem 1

In the empirical illustration, we first test the condition from Theorem 1 based on the RGL curve. Different tests are possible, but we implement a convenient one in which a null hypothesis of $L^A(s) \leq L^B(s)\ \forall s \in [1, 2, ..., l]$ (as in 4) is set against an alternative whereby $\exists s|L^A(s) > L^B(s)$. If we do not reject the null we can state that either A dominates B or the two distributions perfectly overlap, which implies that poverty in A can never be above B for the given choice of weights and deprivation lines, and for any poverty function considered in Theorem 1. Alternatively, rejecting in favour of the alternative means that "A does not dominate B", i.e. either A is dominated by B or the two RGL curves cross (which in turn implies that the poverty comparison is sensitive to the choice of poverty functions, even for a given set of weights and deprivation lines).

In practice, we have a joint intersection null hypothesis: $Ho : L^A(s) = L^B(s)\ \forall s$ against a union alternative $Ha : \exists s|L^A(s) > L^B(s)$. For that purpose, and considering that A and B are independently distributed, we construct the following statistics:

$$T(s) = \frac{L^A(s) - L^B(s)}{\sqrt{\sigma^2_{L^A}(s) + \sigma^2_{L^B}(s)}}, \tag{9}$$

where:

$$\sigma^2_{L^A}(s) \equiv \frac{1}{l^2}\sum_{j=1}^{s}\frac{\sigma^2_{E^A}(l-j+1)}{\pi^A(v_{l-j+1})N^A} \tag{10}$$

Then we test $Ho: T(s) = 0$ against $Ha: T(s) > 0$ for every value of s. Given the requirements of Theorem 1, we conclude that A does not dominates B in terms of Theorem 1 if there is at least one $T(s) > T_\alpha$, where T_α is the right-tail critical value for a one-tailed "z-test" corresponding to a level of significance α. As before, since we test multiple comparisons, the actual size of the whole test is not α. Under reasonable assumptions, it can be shown that it is $\beta = \sum_{i=1}^{l}[l-i+1]\alpha^i(-1)^{i-1}$. With $l = 5$, we choose $\alpha = 0.01$, so that $\beta \approx 0.05$.

3.3 Test of Theorem 3

We follow the same procedure as in the test for Theorem 1, but we note that the RGL curve is constructed differently. Let $\theta_{v_i,v_j} \equiv m_{v_i|v_j}\pi(v_j)$ be the joint probability of attaining a score of v_j in period $t-1$ and a score of v_i in period t. Then:

$$E[c_n^t|v_j]\pi(j) = \sum_{i=1}^{l}\theta_{v_i,v_j}v_i, \tag{11}$$

and:

$$L_\pi(s) = \sum_{j=1}^{s}\sum_{i=1}^{l}\theta_{v_i,v_{l-j+1}}v_i, \quad s = 1,2,...,l. \tag{12}$$

Then we construct a statistic analogous to 9, but $\sigma^2_{L^A}(s)$ (same for B) is replaced by

$$\sigma^2_{L^A_\pi}(s) \equiv \frac{1}{N^A}[\sum_{j=1}^{s}\sum_{i=1}^{l}[\theta_{v_i,v_{l-j+1}}(1-\theta_{v_i,v_{l-j+1}})]v_i^2 - \sum_{j=1}^{s}\sum_{i=1}^{l}\theta_{v_i,v_{l-j+1}}v_i(\sum_{k=1}^{s}\sum_{r=1}^{l}\theta_{v_r,v_{l-k+1}}v_r - \theta_{v_i,v_{l-j+1}}v_i)] \tag{13}$$

The reminder of the test proceeds exactly as in the test for Theorem 1.

3.4 Test of Proposition 2

We recall the definition $F_{v_q|v_j} \equiv \sum_{i=1}^{q} m_{v_i|v_j}$ from the proof for Proposition 2. Because this is a sufficient condition we can set the null hypothesis to be $Ho: F^A_{v_q|v_j} = F^B_{v_q|v_j}\ \forall j \in [1,2,...,l] \wedge q \in [1,2,...,l-1]$ against a union alternative $Ha: \exists(q,j)|F^A_{v_q|v_j} < F^B_{v_q|v_j}$.

The statistic is:

$$T(q,j) = \frac{F^A_{v_q|v_j} - F^B_{v_q|v_j}}{\sqrt{\sigma^2_{F^A}(q,j) + \sigma^2_{F^B}(q,j)}}, \tag{14}$$

where:

$$\sigma^2_{F^A}(q,j) \equiv \frac{F^A_{v_q|v_j}(1 - F^A_{v_q|v_j})}{\pi^A(v_j)N^A} \tag{15}$$

Then we test $Ho : T(q,j) = 0$ against $Ha : T(q,j) < 0$ for every pairwise value (q,j). Given the requirements of Proposition 2, we conclude that A does not dominates B in terms of proposition 2 if there is at least one pair (q,r) such that: $T(q,r) < -T_\alpha$, where T_α is the right-tail critical value for a one-tailed "z-test" corresponding to a level of significance α. As before, since we test multiple comparisons ($l(l-1)$, in fact), the actual size of the whole test is not α. Under reasonable assumptions, it can be shown that it is $\beta = \sum_{i=1}^{l(l-1)}[l(l-1) - i + 1]\alpha^i(-1)^{i-1}$. With $l(l-1) = 20$, we choose $\alpha = 0.001$, so that $\beta \approx 0.02$.

4 Empirical illustration: Multidimensional poverty in Peru

4.1 Background and data

As mentioned, Peru experienced a commodity boom between 2003 and 2007, which translated into high GDP growth rates, from 4 % in 2003 to 8.9 % in 2007, and a steady decrease in monetary poverty headcounts, from 58.7 % in 2004 to 42.4 % in 2007. However, between 2008 and 2013, Peru's economic performance was affected by the world's economic situation: GDP growth fell from 9.8 % in 2008 to 0.9 % in 2009, and then stabilizing around 7 % between 2010 and 2012. Notwithstanding this fluctuation, monetary poverty levels kept decreasing steadily, from 37.3 to 27.8 %. How did the Peruvian population fare in terms of non-monetary multidimensional poverty?

We use Peruvian National Household Surveys (ENAHO). For the non-anonymous assessment, we exploit ENAHO's two recent household panel surveys, spanning 2002-2006 and 2007-2010, each providing 1,570 and 2,260 households, respectively.

Our multidimensional poverty measure relies on four dimensions, and on the household as the unit of analysis. Firstly, household education, comprising two indicators: (1) school delay, which is equal to one if there is a household member in school age who is delayed by at least one year, and (2) incomplete adult primary, which is equal to one if the household head or his/her partner has not completed primary education. The household is considered deprived in education if any of these indicators takes the value of one.

The second dimension considers two indicators on infrastructure dwelling conditions: (i) overcrowding, which takes the value of one if the ratio of the number of household members to the number of rooms in the house is larger than three; and (ii) inadequate construction materials, which takes the value of one if the walls are made of straw or other (almost certainly inferior) material, if the walls are made of stone and mud or wood combined with soil floor, or if the house was constructed at an improvised location inadequate for human inhabitation. The household is deprived in living conditions if any of the above indicators takes the value of one.

The third dimension is access to services. The household is deemed deprived in this dimension if any of the following indicators takes the value of one: (i) lack of electricity for lighting, (ii) lack of access to piped water, (iii) lack of access to sewage or septic tank, and (iv) lack of access to a telephone landline. The fourth dimension is household vulnerability to dependency burdens. The household is deprived or vulnerable if household members who are younger than 14 or older than 64 are three times or more as numerous as those members who are between 14 and 64 years old (i.e. in working age).

We weigh each dimension equally. Therefore the household score can take only any of the following five values: $(0, 0.25, 0.5, 0.75, 1)$.

4.2 Results

Tables 1 through 4 show the transition matrices for deprivation scores. In each matrix the row π shows the initial distribution of scores, the row $E[c_n^t|j]$ shows the expected deprivation score conditional on a score value of j in the initial year, and the row $\hat{\pi}$ shows the ergodic, equilibrium distribution. Overall, all matrices are monotone, therefore the expected deprivation scores increase with the value of the initial, conditioning score. The matrices also exhibit relatively high levels of path-dependence (likelihood of replicating initial conditions in after the transition) as measured by Shorrock's trace index (where 0 means complete immobility and 1 means equality of conditional distributions). The respective values in chronological order are: 0.41, 0.41, 0.33, 0.37. For instance, the probability of being non-deprived in any dimension conditional on having that initial status remains fairly stable, across the matrices, between 82% (2) and 88% (3). Whereas the probability of being deprived in every dimension conditional on having that initial status fluctuates between 42% (2) and 70% (4).

[Table 1 about here.]

[Table 2 about here.]

[Table 3 about here.]

[Table 4 about here.]

4.3 Theorem 1

Table 5 provides the main findings for Theorem 1. It features the vertical coordinates of the RGL curves of expected deprivation scores (the five horizontal coordinates correspond to the number of expected deprivation scores and are common to the four transition matrices). The ensuing partial ordering, related to Theorem 1, states that if the initial relative frequencies of scores were identical within and across samples, then the ex-ante expected social poverty induced by matrix 1 (2002-2004) would be lower than the levels produced by matrices 3 (2007-2008) and 4 (2008-2010), for any social poverty function that increases both with higher conditional expected deprivation scores and with higher inequality between them (by contrast the point estimates of the RGL curves of matrix 1 and matrix 2 (2004-2006) cross between the horizontal coordinates 1 and 2). Likewise, matrix 2 induces lower ex-ante expected poverty than matrices 3 and 4. Finally, a similar robust ordering cannot be established between matrices 3 and 4 since their two respective RGL curves cross (between horizontal coordinates 1 and 2). In summary, the conditional distributions of expected deprivation scores produced by the pre-crisis transition matrices second-order dominate the distributions yielded by the crisis/post-crisis matrices, according to theorem 1.

[Table 5 about here.]

Table 6 shows the test statistics for the tests of Theorem 1. The tests confirm the dominance results featured in Table 5. Additionally, the only instance in which the RGL curve of matrix 1 (2002-2004) is above that of matrix 2 (2004-2006) is not statistically significant ($T(1) = 0.778$). Hence, actually we cannot reject the hypothesis that the RGL curve of matrix 1 is never above that of matrix 2, which means that either both curves perfectly overlap, or the former RGL curve dominates the latter. Likewise, the only instance in which the RGL curve of matrix 3 (2007-2008) is below that of matrix 4 (2008-2010) is not statistically significant ($T(1) = -1.484$). With these results we cannot reject the hypothesis that the RGL curve of matrix 3 is never below that of matrix 4. Thus these statistical results provide us with a complete ordering according to Theorem 1, whereby if the initial relative frequencies of scores were identical within and across samples, then the ex-ante expected social poverty induced by matrix 1 (2002-2004) would be lower than the levels produced by all the other matrices, for any social poverty function that increases both with higher conditional expected deprivation scores and with higher inequality between them. Likewise the expected social poverty induced by matrix 2 (2004-2006) would be lower than the levels produced by the two subsequent matrices. Finally, matrix 3 (2007-2008) would yield the highest expected social poverty among all matrices, in terms of the aforementioned criteria.

[Table 6 about here.]

4.4 Theorem 2

Table 7 provides the main findings for Theorem 2. It shows again the values of all the expected deprivation scores, conditioned by the different values of $c_n^{t-1}$for each transition matrix. Vector dominance is absent from every possible pairwise comparison between transition periods. Therefore we conclude that the transition matrices cannot be ordered according to Theorem 2, which is based on equal Π between compared samples. This means that the social expected poverty rankings, in every pairwise comparison, depend on the choice of Π, i.e. no ordering is robust. For example, as it turns out, each matrix could induce the lowest ex-ante expected poverty depending on Π. Should $\pi(1) \to 1$ then matrix 2 (2004-2006) would induce the lowest poverty (while inducing the highest level if $\pi(0) \to 1$). If $\pi(0.25) \to 1$ then matrix 1 (2002-2004) achieves the lowest level, while if $\pi(0.5) \to 1$ then matrix 4 (2008-2010) does. Finally, if $\pi(0) \to 1$ then matrix 3 (2007-2008) reaches the lowest level.

[Table 7 about here.]

Table 8 shows the test statistics for the tests of Theorem 2. We reject the null hypothesis that at least one expected deprivation score from matrix 1 (2002-2004) is significantly higher than its counterpart in matrix 2 (2004-2006). In fact the evidence suggests dominance of matrix 1 in terms of Theorem 2. Likewise we find dominance of matrix 1 over matrix 3 (2007-2008), as well as dominance of matrix 2 over matrix 4 (2008-2010), and dominance of matrix matrix 4 over matrix 3. Finally, we cannot ascertain dominance between matrix 1 and matrix 4 since the pairwise differences in their expected deprivation scores are all statistically insignificant.

[Table 8 about here.]

4.5 Theorem 3 with initial distributions

Table 9 provides the main findings for Theorem 3, using actual initial distributions, Π. It features the vertical coordinates of the RGL curves, as in Table 5. The ensuing partial ordering, related to Theorem 3, states that, with the actual initial relative frequencies of scores, the ex-ante expected social poverty induced by matrix 4 (2008-2010) would be lower than the levels produced by matrix 3 (2007-2008), for any social poverty function that increases both with higher conditional expected deprivation scores and with higher inequality between them. All the other comparisons feature RGL curve-crossing.

[Table 9 about here.]

Table 10 shows the test statistics for the tests of Theorem 3 with actual initial distributions, Π. The evidence shows matrix 1 (2002-2004) being dominated by both matrix 3 (2007-2008) and matrix 4 (2008-2010) (since the curve-crossings are not statistically significant).Likewise we find that matrix 4 dominates matrix 2 (2004-2006). In the case of the other three comparisons, the pairwise differences between the RGL curves are not statistically significant at all, i.e. in these cases we would not be able to reject a null hypothesis of perfect curve overlap.

[Table 10 about here.]

4.6 Theorem 3 with ergodic distributions

Table 11 provides the main findings for Theorem 3, using ergodic distributions, $\hat{\Pi}$. It features the vertical coordinates of the RGL curves, as in previous tables. The ensuing ordering, related to Theorem 3, states that, with the ergodic distributions of scores, the ex-ante expected social poverty induced by matrices 1 (2002-2004) and 4 (2008-2010) would be lower than the levels produced by the other two matrices, for any social poverty function that increases both with higher conditional expected deprivation scores and with higher inequality between them. Meanwhile, a robust ordering between these two matrices is not possible. Additionally, matrix 2 (2004-2006) induces lower ex-ante expected poverty than matrix 3 (2007-2008).

[Table 11 about here.]

Tables 12, 13, 14, 15, 16, and 17 show the test statistics for the tests of Proposition 2. The results are coherent with the point-estimate comparisons based on Table 11. Even though Proposition 2 requires consistency across the signs of all the statistics, when we test, it turns out that the statistics with the "wrong" sign (in terms of being inconsistent with the comparisons in Table 11) are not statistically significant individually. In fact most statistics do not appear statistically significant at the stringent level of significance set above ($\alpha = 0.001$ for each statistic individually, in order to achieve an overall level of $\beta = 0.02$). However, even if the statistics with the "wrong" sign were significant and they appear along with statistics with the "right" sign for a given hypothesis, we must recall that Proposition 2 is sufficient, but not necessary, in order to ensure the fulfillment of a robust ordering in the spirit of Theorem 11.

[Table 12 about here.]

[Table 13 about here.]

[Table 14 about here.]

[Table 15 about here.]

[Table 16 about here.]

[Table 17 about here.]

5 Concluding remarks

The non-anonymous approach to robust inter-temporal poverty comparisons with counting measures, presented in this paper, is basically an intra-generational mobility assessment comparing the outcomes of the same people across different periods. Generally, when non-anonymous conditions are fulfilled, we can conclude that, *ex-ante*, the distribution of expected deprivation scores (conditioned on different initial deprivation scores) of A second-order dominates that of B; meaning, *inter alia*, that the distribution in A features both lower average expected deprivation scores, and less dispersion than B's in Lorenz-consistency terms. In that sense the poverty reduction experience is more “pro-poorest” in A than in B.

More specifically we proposed different conditions based on alternative ways of constructing the distributions of expected deprivation scores. The first two methods simply compare the vectors of expected deprivation scores without any information on their actual distributions (or expected equilibrium distributions in the future). But they differ in their welfare interpretation. In particular, when Theorem 2 holds, the social poverty comparison is robust to all possible distributions of initial (conditioning) scores, as long as the distributions are the same between the compared samples. By contrast, the latter two methods compare distributions of expected deprivation scores so that their respective Reversed Generalized Lorenz (RGL) curves depend both on the vector of expected deprivation scores and their distribution. In one method the distribution is taken to be the initial distribution of scores. In the other case the distribution is the ergodic one. Of course, all these conditions hold only for specific choices of deprivation lines and dimensional weights. With alternative selections, the conditions must be tested again.

Our empirical illustration of the non-anonymous condition, using the Peruvian panel datasets, showed that the existence and nature of robust partial orderings depend crucially on the theorem being applied. For instance, in the case of Theorem 1 the mobility matrix of deprivation scores corresponding to the 2002-2004 period induced a preferable distribution of expected deprivation scores vis-a-vis all the other matrices, i.e. those of 2004-2006, 2007-2008, and 2008-2010. The second-best matrix was that of period 2004-2006, which was also preferable to the two crisis/post-crisis matrices (but not to its predecessor, 2002-2004). At the other end, the matrix for 2007-2008 turned out the least desirable according to Theorem 1.

In the case of Theorem 2, it was not possible to rank the mobility matrices robustly just by comparing the point estimates of the expected deprivation scores. However, since many of the pairwise differences between scores of different matrices were not statistically significant, we can conclude that the matrix of 2002-2004 dominates those of 2004-2006 and 2007-2008 (whereas no difference is statistically significant in the comparison of scores between 2002-2004 and 2008-2010). We also found that matrix 2004-2006 is dominated by all the others, and that the matrix of 2008-2010 dominates that of 2007-2008.

When we made the RGL curves depend not only on the vectors of expected deprivation scores, but also on their distributions in each society, we found few statistically significant comparisons for the chosen levels of significance, in the case of Theorem

3 and initial distributions of conditioning scores. But we did find that the matrix of 2008-2010 dominates that of 2002-2004. If we allow for slightly higher test size, the dominance of 2007-2008 over 2002-2004, and of 2008-2010 over 2004-2006, also turn out statistically significant. This result stands in contrast to that for Theorem 1, in which the mobility matrix of 2002-2004 was the most preferable, followed by 2004-2006. The driving source of difference must be the initial distribution, which is first-order stochastically dominated by those of the matrices from later transition periods.

A somewhat intermediate result is provided by the case of Theorem 3 with ergodic distributions. Since the ergodic distribution of 2002-2004 is more preferable than the other matrices', which is not the case when comparing their respective initial distributions, it turns out that, again, the matrix of 2002-2004 robustly induces lower social expected poverty than all the other matrices, except for 2008-2010; which, in turn, also dominates 2004-2006 and 2007-2008. This latter criterion cannot order 2002-2004 and 2008-2010, since their respective RGL curves based on ergodic distributions cross.

References

Alkire, S. and J. Foster (2011). Counting and multidimensional poverty measurement. *Journal of Public Economics 95*(7-8), 476–87.

Alkire, S. and S. Seth (2014). Measuring and decomposing inequality among the multidimensionally poor using ordinal data: A counting approach. OPHI Working Paper 68.

Ben Haj Kacem, R. (2013). Monetary versus non-monetary pro-poor growth: Evidence from rural ethiopia between 2004 and 2009. *Economics: The Open-Access, Open-Assessment E-Journal, 7*(26).

Benabou, R. and E. Ok (2001). Mobility as progressivity: ranking income processes according to equality of opportunity. Working Paper 8431 NBER.

Berenger, V. and F. Bresson (2012). On the "pro-poorness" of growth in a multidimensional context. *Review of Income and Wealth 58*(3), 457–80.

Boccanfuso, D., J. Bosko Ki, and C. Menard (2009). Pro-poor growth measurements in a multidimensional model: A comparative approach. GREDI Working Paper 09-22.

REFERENCES

Chakravarty, S. and C. D'Ambrosio (2006). The measurement of social exclusion. *Review of Income and wealth 52*(3), 377–98.

Chakravarty, S. and C. Zoli (2009). Social exclusion orderings. Working Papers Series. Department of Economics. University of Verona.

Dardanoni, V. (1995). Income distribution dynamics: monotone markov chains make light work. *Social Choice and Welfare 12*, 181–92.

Fleurbaey, M. and V. Peragine (2013). Ex ante versus ex post equality of opportunity. *Economica 80*(317), 118–30.

Formby, J., J. Smith, and B. Zheng (2004). Mobility measurement, transition matrices and statistical inference. *Journal of Econometrics 120*, 181–205.

Gallegos, J. and G. Yalonetzky (2014). Robust pro-poorest poverty reduction with counting measures: the anonymous case. mimeograph.

Grimm, M. (2007). Removing the anonymity axiom in assessing pro-poor growth. *Journal of Economic Inequality 5*, 179–97.

Kremer, M., A. Onatski, and J. Stock (2001). Searching for prosperity. *Carnegie-Rochester Conference Series on Public Policy 55*, 275–303.

Lasso de la Vega, C. (2010). Counting poverty orderings and deprivation curves. In J. Bishop (Ed.), *Research on Economic Inequality*, Volume 18, Chapter 7, pp. 153–72. Emerald.

Marshall, A., I. Olkin, and B. Arnold (2011). *[I]nequalities: [T]heory of [M]ajorization and its applications* (Second ed.). Springer Series in Statistics. Springer.

Shorrocks, A. (1983). Ranking income distributions. *Economica 50*, 3–17.

Silber, J. and G. Yalonetzky (2013). *Poverty and Social Exclusion. New methods of analysis*, Chapter 2, pp. 9–37. Routledge Advances in Social Economics. Routledge.

TABLES

Table 1: Transition matrix of deprivation scores, Peru, 2002-2004

		2002				
		0	0.25	0.5	0.75	1
	0	0.87	0.21	0.02	0.0	0.0
	0.25	0.11	0.65	0.20	0.04	0.0
2004	0.5	0.02	0.14	0.67	0.33	0.09
	0.75	0.0	0.0	0.11	0.61	0.36
	1	0.0	0.0	0.0	0.02	0.55
	π	0.18	0.28	0.36	0.17	0.01
	$E[c_n^t \mid j]$	0.039	0.235	0.467	0.653	0.864
	$\hat{\pi}$	0.47	0.27	0.20	0.06	0.00

Table 2: Transition matrix of deprivation scores, Peru, 2004-2006

		2004				
		0	0.25	0.5	0.75	1
	0	0.82	0.19	0.02	0.0	0.0
	0.25	0.15	0.68	0.19	0.03	0.0
2006	0.5	0.03	0.12	0.69	0.21	0.0
	0.75	0.0	0.01	0.11	0.74	0.58
	1	0.0	0.0	0.0	0.02	0.42
	π	0.23	0.28	0.34	0.15	0.01
	$E[c_n^t\|j]$	0.051	0.239	0.473	0.686	0.854
	$\hat{\pi}$	0.35	0.31	0.23	0.11	0.00

Table 3: Transition matrix of deprivation scores, Peru, 2007-2008

		2007				
		0	0.25	0.5	0.75	1
	0	0.88	0.14	0.01	0.0	0.0
	0.25	0.10	0.71	0.15	0.01	0.0
2008	0.5	0.02	0.14	0.74	0.24	0.0
	0.75	0.0	0.0	0.10	0.74	0.40
	1	0.0	0.0	0.01	0.01	0.60
	π	0.27	0.27	0.29	0.15	0.01
	$E[c_n^t\|j]$	0.034	0.251	0.489	0.686	0.900
	$\hat{\pi}$	0.33	0.26	0.28	0.12	0.01

TABLES

Table 4: Transition matrix of deprivation scores, Peru, 2008-2010

		2008				
		0	0.25	0.5	0.75	1
	0	0.86	0.17	0.03	0.0	0.0
	0.25	0.13	0.68	0.24	0.05	0.0
2010	0.5	0.01	0.14	0.64	0.26	0.05
	0.75	0.0	0.01	0.10	0.67	0.25
	1	0.0	0.0	0.0	0.02	0.70
	π	0.283	0.267	0.293	0.148	0.01
	$E[c_n^t \| j]$	0.039	0.250	0.451	0.666	0.913
	$\hat{\pi}$	0.43	0.32	0.18	0.06	0.00

Table 5: RGL curves of expected deprivation scores (as defined in 4). Vertical coordinates.

	1	2	3	4	5
2004|2002	0.863	1.516	1.983	2.218	2.257
2006|2004	0.854	1.540	2.013	2.252	2.304
2008|2007	0.900	1.586	2.074	2.325	2.359
2010|2008	0.913	1.578	2.029	2.279	2.318

TABLES

Table 6: Z-statistics for tests of Theorem 1.

$Ho: L^A(s) = L^B(s)\ \forall s$ $Ha: \exists s \vert L^A(s) > L^B(s)$	1	2	3	4	5
A = 2004|2002; B = 2006|2004	0.778	-1.902	-2.367	-2.655	-3.610
A = 2004|2002; B = 2008|2007	-3.217	-6.046	-7.804	-9.037	-8.615
A = 2004|2002; B = 2010|2008	-4.149	-5.166	-3.739	-4.897	-4.881
A = 2006|2004; B = 2008|2007	-5.103	-4.972	-6.496	-7.588	-5.765
A = 2006|2004; B = 2010|2008	-6.094	-3.896	-1.563	-2.588	-1.360
A = 2008|2007; B = 2010|2008	-1.484	0.866	5.140	5.146	4.620

Table 7: Conditional expected deprivation scores

Transition c_n^{t-1}	0	0.25	0.5	0.75	1
2004|2002	0.039	0.235	0.467	0.652	0.864
2006|2004	0.051	0.239	0.473	0.686	0.854
2008|2007	0.034	0.251	0.489	0.686	0.900
2010|2008	0.039	0.250	0.451	0.666	0.913

Table 8: Z-statistics for tests of Theorem 2.

$Ho: E^A[c_n^t \mid v_j] = E^B[c_n^t \mid v_j]\ \forall j$ $Ha: \exists j \mid E^A[c_n^t \mid v_j] > E^B[c_n^t \mid v_j]$	1	2	3	4	5
A = 2004|2002; B = 2006|2004	-1.427	-0.394	-0.651	-2.578	0.155
A = 2004|2002; B = 2008|2007	0.562	-1.708	-2.485	-2.942	-0.643
A = 2004|2002; B = 2010|2008	-0.030	-1.544	1.825	-1.077	-0.829
A = 2006|2004; B = 2008|2007	2.253	-1.279	-1.764	-0.015	-1.020
A = 2006|2004; B = 2010|2008	1.673	-1.132	2.491	1.634	-1.218
A = 2008|2007; B = 2010|2008	-0.780	0.100	4.590	1.905	-0.296

Table 9: RGL curves of expected deprivation scores (as defined in 5, using initial distributions). Vertical coordinates.

	1	2	3	4	5
2004|2002	0.006	0.120	0.287	0.351	0.358
2006|2004	0.007	0.109	0.270	0.336	0.348
2008|2007	0.008	0.114	0.255	0.324	0.333
2010|2008	0.008	0.107	0.239	0.306	0.317

TABLES

Table 10: Z-statistics for tests of Theorem 3.

$Ho: L_\pi^A(s) = L_\pi^B(s)\ \forall s$ $Ha: \exists s \vert L_\pi^A(s) > L_\pi^B(s)$	1	2	3	4	5
$A = 2004\vert 2002$; $B = 2006\vert 2004$	-0.169	0.994	1.021	0.885	0.627
$A = 2004\vert 2002$; $B = 2008\vert 2007$	-0.611	0.559	2.101	1.725	1.566
$A = 2004\vert 2002$; $B = 2010\vert 2008$	-0.649	1.272	3.239	2.874	2.604
$A = 2006\vert 2004$; $B = 2008\vert 2007$	-0.422	-0.528	0.991	0.761	0.883
$A = 2006\vert 2004$; $B = 2010\vert 2008$	-0.460	0.190	2.136	1.916	1.923
$A = 2008\vert 2007$; $B = 2010\vert 2008$	-0.041	0.801	1.283	1.292	1.166

Table 11: RGL curves of expected deprivation scores (as defined in 4, using ergodic distributions). Vertical coordinates.

	1	2	3	4	5
2004|2002	0.003	0.044	0.142	0.205	0.223
2006|2004	0.004	0.084	0.191	0.265	0.282
2008|2007	0.007	0.090	0.223	0.289	0.301
2010|2008	0.004	0.053	0.140	0.218	0.234

Table 12: Z-statistics for tests of Proposition 2, $Ho : F^{2002-2004}_{v_q|v_j} = F^{2004-2006}_{v_q|v_j} \ \forall(q,j)$, $Ha : \exists(q,j)|F^{2002-2004}_{v_q|v_j} < F^{2004-2006}_{v_q|v_j}$

	0	0.25	0.5	0.75	1
0	1.579	0.847	0.560	-	-
0.25	0.399	-0.591	0.601	0.627	-
0.5	0.184	1.130	0.166	2.984	1.048
0.75	0.184	-	0.661	-0.090	-0.622

Table 13: Z-statistics for tests of Proposition 2, $Ho : F^{2002-2004}_{v_q|v_j} = F^{2007-2008}_{v_q|v_j}\ \forall(q, j)$**,** $Ha : \exists(q, j)|F^{2002-2004}_{v_q|v_j} < F^{2007-2008}_{v_q|v_j}$

	0	0.25	0.5	0.75	1
0	-0.610	2.757	2.289	-1.008	-
0.25	-0.373	-0.058	2.751	2.227	-
0.5	0.964	-0.856	0.066	3.080	1.048
0.75	-	-	1.522	-1.070	0.293

Table 14: Z-statistics for tests of Proposition 2, $Ho: F^{2002-2004}_{v_q|v_j} = F^{2008-2010}_{v_q|v_j}\ \forall(q,j)$**,** $Ha: \exists(q,j)|F^{2002-2004}_{v_q|v_j} < F^{2008-2010}_{v_q|v_j}$

	0	0.25	0.5	0.75	1
0	0.334	1.621	-0.870	-	-
0.25	-0.761	0.454	-1.902	-0.359	-
0.5	-	1.463	-0.533	1.462	0.411
0.75	-	-	-0.986	0.165	0.849

Table 15: Z-statistics for tests of Proposition 2, $Ho : F^{2004-2006}_{v_q|v_j} = F^{2007-2008}_{v_q|v_j} \ \forall(q,j)$**,** $Ha : \exists(q,j)|F^{2004-2006}_{v_q|v_j} < F^{2007-2008}_{v_q|v_j}$

	0	0.25	0.5	0.75	1
0	-2.476	1.852	1.735	-1.008	-
0.25	-0.856	0.589	2.092	1.519	-
0.5	0.850	-1.860	-0.108	-0.192	-
0.75	-0.552	-	0.839	-0.921	1.020

Table 16: Z-statistics for tests of Proposition 2, $Ho: F^{2004-2006}_{v_q|v_j} = F^{2008-2010}_{v_q|v_j}\ \forall(q,j)$**,** $Ha: \exists(q,j)|F^{2004-2006}_{v_q|v_j} < F^{2008-2010}_{v_q|v_j}$

	0	0.25	0.5	0.75	1
0	-1.512	0.709	-1.455	-	-
0.25	-1.260	1.102	-2.518	-1.016	-
0.5	-0.542	0.184	-0.700	-1.730	-1.025
0.75	-0.542	-	-1.439	0.254	1.615

Table 17: Z-statistics for tests of Proposition 2, $Ho: F^{2007-2008}_{v_q|v_j} = F^{2008-2010}_{v_q|v_j}\ \forall(q,j)$**,** $Ha: \exists(q,j)|F^{2007-2008}_{v_q|v_j} < F^{2008-2010}_{v_q|v_j}$

	0	0.25	0.5	0.75	1
0	1.211	-1.277	-3.296	1.008	-
0.25	-.524	0.568	-4.924	-2.797	-
0.5	-1.063	2.360	-0.631	1.719	-1.025
0.75	-	-	-2.252	1.335	0.666