

Towards a New Definition of Shared Prosperity: A Dynamic Perspective from Three Countries

Peter Lanjouw (VU University Amsterdam)
with Hai-Anh Dang (World Bank)

International Workshop
“Inequality in a Recessionary World: From Global to Local”

University of Valencia
June 19, 2015

Introduction

- Global poverty reduction in recent decades has been impressive
 - Between 1990 and 2011 global extreme poverty has halved to roughly 14.5% of world population
 - based on \$1.25 PPP global poverty line
- Large parts of the world no longer register significant poverty rates at this austere line
- As extreme poverty recedes, countries are increasingly asking about the *quality* of their growth as a yardstick for progress
 - Are the poor sharing in rising prosperity?

Introduction

- The World Bank has proposed tracking *shared prosperity*
 - Growth in average income of the bottom 40%
 - simple to calculate and explain
 - country level indicator
 - unbounded
 - survey-based
 - readily linked to inequality
 - World Bank (2014) provides overview

Elements of an alternative (complementary) approach

- At the country level, poverty lines are commonly established on the basis of national criteria
 - These national poverty lines often lie above global line
 - With rising incomes national poverty lines typically “drift” upward (“weakly” relative lines)
- Proposal: Link discussion of rising prosperity to declining national poverty
- BUT may also want to extend notion of prosperity beyond poverty to population segments above poverty line

Looking beyond declining poverty

- It is possible that significant percentage of population above the poverty line still faces a significant *risk* of falling back into poverty
 - “Vulnerable”: non-poor who face a heightened risk of falling into poverty
 - “Secure” (or “middle class”): neither poor nor vulnerable
- Proposal: consider rising shared prosperity in terms of growing “middle class” and declining (national) poverty and vulnerability
 - Some precedence for pursuing this route (eg. WB reports on middle class in Latin America; media articles on “fragile middle class”)
 - Existing methods employ somewhat ad-hoc definitions of “middle class”

Challenges

- Definition of vulnerability (and middle class) proposed here depends on analysis of poverty dynamics
 - Estimating likelihood of falling into poverty
 - Straightforward with panel data
- In developing countries panel data are scarce
 - Rarely nationally representative
 - Prone to attrition; measurement error
- This paper employs “synthetic” panel methods to circumvent data constraints
 - Estimate poverty transitions to identify vulnerable population
 - Dang, Lanjouw, Luoto, McKenzie (2014) – bounds approach
 - Dang and Lanjouw (2013) – parametric approach yielding point estimates

Outline

- Measuring vulnerability
- Application to U.S. and Vietnam
- Review of synthetic panel methodology
 - Validation exercise with Vietnamese data
- Application to India data
- Conclusions

Vulnerability

- Link vulnerability to the notion of susceptibility to something harmful that has not yet occurred.
 - Distinguish the vulnerable from the poor
 - Pritchett et al (2000) specify a vulnerability line as the level of income below which a household experiences a greater than even chance of experiencing poverty in near future.
 - “vulnerable” by this definition includes the currently poor.
- Set a vulnerability line that builds on the poverty line and separates the population into three groups:
 - Poor, vulnerable, secure

Two approaches

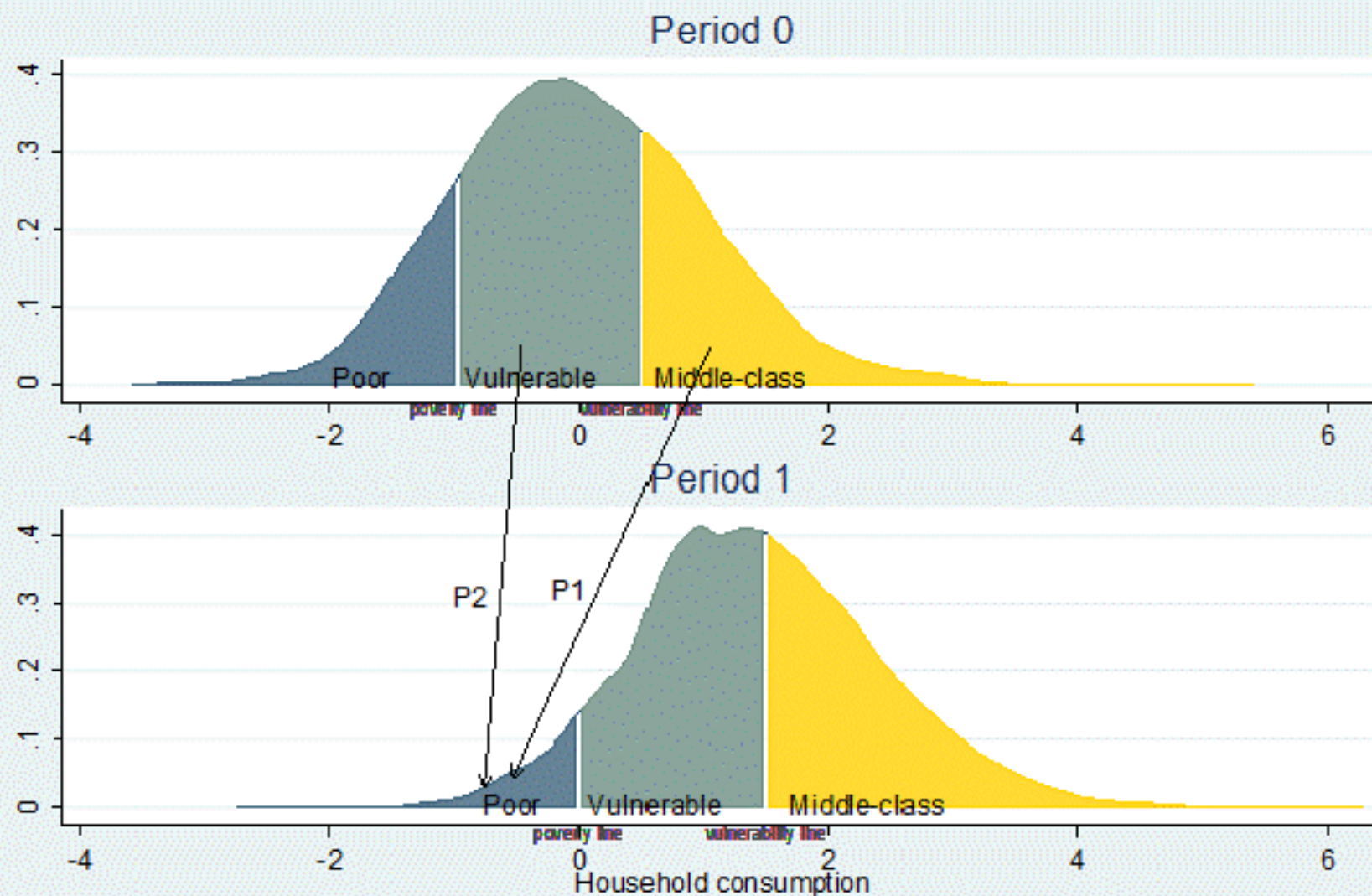
Approach 1

- Specify the highest acceptable level of risk of falling into poverty; more is considered excessive/unacceptable.
- Identify population whose risk of falling into poverty in next period is at this level *or lower*.
- Define vulnerability line as the lower bound income level of this population (ie, the middle class).
- Population that lies between the poverty line and this vulnerability line is designated as vulnerable.

Two approaches

Approach 2

- Specify the highest acceptable level of risk of falling into poverty, more than that is considered excessive/unacceptable.
- Identify population that is clearly not poor, but whose risk of falling into poverty in the next period is at this level *or higher*.
- Define the vulnerability line as the upper bound income level for this population (ie, the vulnerable).



Note
 Household consumption distribution in each period is divided into 3 groups: poor, vulnerable, and middle-class, which are shaded differently. P1 and P2 respectively represents the insecurity index (percentage of the population that are in the middle class group in period 0 but in the poor group in period 1) and the vulnerability index (percentage of the population that are in the vulnerable group in period 0 but in the poor group in period 1)

Features

- Specifying a common cut-off risk level allows comparisons of vulnerability even across settings where income levels and poverty lines vary.
- Estimation of these vulnerability lines is simple
 - “light” demands of underlying panel data
 - Can be straightforwardly implemented also with synthetic panels.
 - Can be contrasted with other proposed approaches
 - Chaudhuri (2003) uses cross-section data but makes restrictive assumptions
 - Lopez-Calva and Ortiz-Juarez (2014) rely on panel data and make a number of parametric assumptions
- Note: The approach is intended to identify population groups rather than establish vulnerability of individual households.

Implementation

- Approach 1

estimate:
$$p^1 = \frac{P(y_1 \leq Z_1 \cap y_0 > V_0)}{P(y_0 > V_0)}$$

- Approach 2

estimate:
$$p^2 = \frac{P(y_1 \leq Z_1 \cap Z_0 < y_0 < V_0)}{P(Z_0 < y_0 < V_0)}$$

- Solve empirically for V_0

Data

- US: panel income data from PSID for 2005, 2007 and 2009 (reference to previous year)
 - 5,335 panel households
- Vietnam: panel consumption data from VHLSS for 2004, 2006, 2008
 - 1,800 panel households for 2004, 2006 and 2008
 - 3,735 panel households for 2006 and 2008

Approach 1 in US and Vietnam

No	Vulnerability index	United States			Vietnam		
		Vulnerability line	Increase (%)	Pop. share with consumption above V-line	Vulnerability line	Increase (%)	Pop. share with consumption above V-line
1	6	N/A	N/A	N/A	2560	0	34.2
2	5	N/A	N/A	N/A	2800	9	30.4
3	4	13305	0	89.0	3080	20	75.7
4	3	17905	35	83.5	3320	30	71.6
5	2	26505	99	73.6	3920	53	61.0
6	1	61305	361	40.8	5320	108	40.7

Note: Vulnerability lines are in \$US per household and D'000 per capita respectively for the US and Vietnam. The relative increases of the vulnerability line from the poverty line for each country is shown under the columns "Increase" (columns 4 and 7). All numbers are estimated with true panel data and weighted with population weights. Estimation sample sizes are 5,335 panel households for the US, and 3,735 panel households for Vietnam. The incremental values for iteration are \$US100 and D20,000 respectively for the US and Vietnam. The exchange rate is US\$1 for D16,302 in 2008 (World Bank, 2013).

Approach 2 in US and Vietnam

No	Vulnerability index	United States			Vietnam		
		Vulnerability line	Increase (%)	Pop. share with consumption above poverty line but less than Y-line	Vulnerability line	Increase (%)	Pop. share with consumption above poverty line but less than Y-line
1	22	N/A	N/A	N/A	3340	30	12.9
2	21	N/A	N/A	N/A	3380	32	13.3
3	20	N/A	N/A	N/A	3500	37	15.3
4	19	21100	59	9.0	3600	43	18.2
5	18	22200	67	10.3	3700	45	18.0
6	17	22800	71	11.0	3780	48	20.6
7	16	24300	83	12.6	3920	53	23.3
8	15	25100	89	13.9	4140	62	26.7
9	14	27900	110	16.8	4260	66	28.6
10	13	29000	118	18.5	4500	76	32.2
11	12	30400	129	20.1	4860	80	37.1
12	11	32400	144	22.2	5140	101	40.9
13	10	36900	177	26.4	5480	114	45.5
14	9	45900	245	34.1	6080	138	51.6
15	8	50600	280	39.2	7080	177	59.3
16	7	59100	344	46.1	8980	251	69.6
17	6	74400	459	56.8	17300	384	81.8
18	5	104400	685	69.9	N/A	N/A	N/A
19	4	290900	2086	86.9	N/A	N/A	N/A

Note: Vulnerability lines are in \$US per household and \$1000 per capita respectively for the US and Vietnam. The relative increases of the vulnerability line from the poverty line for each country is shown under the columns 'Increase' (columns 4 and 7). All numbers are estimated with true panel data and weighted with population weights. Estimation sample sizes are 3,333 panel households for the US, and 3,733 panel households for Vietnam. The incremental values for iteration are \$US100 and \$20,000 respectively for the US and Vietnam. The exchange rate is US\$1 for \$16,302 in 2008 (World Bank, 2013).

Vulnerability in US and Vietnam

- Approach 2 with a cut-off risk of 10%
- Estimate vulnerability line in 2006
 - Examine risk of falling into poverty between 2006 and 2008
- Inflate/deflate vulnerability line to 2008 and 2004, respectively using CPI data
- Compare percentage of population poor, vulnerable and “prosperous” in 2004 and 2008

Trends in Vulnerability and Poverty in US and Vietnam: 2004-2008

	US		Vietnam	
	2004	2008	2004	2008
Poor	8.9	10.0	20.0	14.3
Vulnerable	26.4	27.9	51.0	49.7
Middle class	64.7	62.1	29.0	36.0

Application to India

- No panel data available in India
- Employ cross-sectional consumption data from NSS for 2004/5, 2009/10
- Dang, Lanjouw, Luoto and McKenzie (2014), and Dang and Lanjouw (2013) outline a procedure to construct synthetic panels out of cross section data.
- Based on imputation models, reliant on time-invariant correlates of consumption.
- Fairly extensive validation work suggest approach is reasonably reliable.

The proposed approach

(Dang, Lanjouw, Luoto and McKenzie, JDE 2014)

- Combines ideas of poverty-mapping with pseudo-panel ideas.
- Will set out for case of 2 rounds, can be extended easily to multiple rounds.
- Let x_{i1} be characteristics of household i in time period 1, which are observed in both the round 1 and round 2 surveys:
 - All time-invariant characteristics (language, religion, ethnicity)
 - Characteristics of household head if the head doesn't change across rounds (sex, place of birth, parental education, etc.)
 - Can include time-varying characteristics that can easily be recalled for round 1 in round 2
 - E.g. whether household head was employed in round 1, place of residence in round 1, whether household has a TV in round 1, etc.
 - Can also include time invariant ancillary variables (Census, GIS, etc.) at a more aggregated level

Projections

- Project round 1 consumption or income onto x_{i1} :

$$y_{i1} = \beta_1' x_{i1} + \varepsilon_{i1}$$

- Project round 2 consumption or income onto same set of characteristics as they appear at time of second round:

$$y_{i2} = \beta_2' x_{i2} + \varepsilon_{i2}$$

- Then we are interested in knowing quantities such as: $Pr(y_{i1} < p | y_{i2} > p)$

Don't observe for the same household



Proposed method

- *Step one:* Use the sample of households observed in round 1, and regress y_{i1}^1 on x_{i1}^1
 - Obtain the OLS estimator $\hat{\beta}_1$ and the residuals:
$$\hat{\varepsilon}_{i1}^1 = y_{i1}^1 - \hat{\beta}_1' x_{i1}^1$$
 - Superscript 1 denotes that these are observations for households observed in round 1 only.
- *Step two:* For each household observed in round 2, take a random draw with replacement from the empirical distribution of residuals, then combine with parameter estimate and known x to estimate round 1 income or consumption:

$$\hat{y}_{i1}^2 = \hat{\beta}_1' x_{i1}^2 + \tilde{\varepsilon}_{i1}^2$$

Proposed method

- *Step Three:* calculate movements into and out of poverty using \hat{y}_{i1}^2 in place of the unobserved round 1 variable:

$$Pr(\hat{y}_{i1}^2 < p | y_{i2}^2 > p)$$

- *Step Four:* Repeat steps 1-3 R times, and take average of the quantity of interest over the R replications.

Under what conditions will this be consistent?

- *Condition 1:* the underlying population sampled is the same in round 1 and round 2
 - Requires measure of consumption to be same from round to round,
 - Assumption implies that households in period 2 that have similar characteristics to those of households in period 1 would have achieved the same consumption levels in period 1 or vice versa.

Under what conditions will this be consistent?

- *Condition 2:* ε_{i1} is independent of y_{i2} . This requires ε_{i1} to be independent of ε_{i2}

Won't hold if:

- Error term contains individual fixed effect
- If shocks to consumption or income are non-transitory.

- We expect in many cases this condition to be violated.
 - So long as errors positively correlated (which seems likely in most cases), this will *overstate* mobility, providing an *upper bound* on movements into and out of poverty.
 - If errors are negatively correlated then our method wouldn't provide a bound.
 - We don't expect a negative correlation *on average*
 - DLLM (2014) demonstrate this empirically with real panel data.

Lower bound method

- Instead assume the prediction error for household i in round 1 is the same as it is for round 2 (perfect positive autocorrelation).
- *Step One:* for sample of households surveyed in round 2, obtain OLS residuals:

$$\hat{\varepsilon}_{i2}^2 = y_{i2}^2 - \hat{\beta}_2' x_{i2}^2$$

- *Step Two:* then estimate round 1 income or consumption as

$$\tilde{y}_{i1}^2 = \hat{\beta}_1' x_{i1}^2 + \hat{\varepsilon}_{i2}^2$$

- *Step Three:* Use the estimated y from step 2 to calculate poverty dynamic of interest.

Imposing parametric assumptions

- Assume ε_{i1} and ε_{i2} have a bivariate normal distribution
- ρ is the correlation coefficient between these two error terms (assumed positive).
- DLLM bounds assume ρ being equal to its maximum value (1) and minimum value (0)
- But the true value of ρ in all likelihood lies somewhere in between these two values.
- Parametric approach allows us to “plug” in a value of ρ

Imposing Parametric Assumptions

- DLLM (2014) explore existing panel surveys to get a sense of plausible values for ρ
 - Evidence from Indonesia, Vietnam, Chile, Nicaragua and Peru suggest ρ might range between 0.5-0.8
 - Preliminary evidence from SILC data from Europe suggest a range for ρ between 0.3-0.8
- Dang and Lanjouw (2013) propose an approach to estimate ρ directly, based on birth cohort-aggregated consumption correlations.

Validation

- Take panel data for Vietnam for 2006-2008 and treat the two halves as though they were cross section data
- Apply synthetic panel method

Validation of Synthetic Panel Approach: Vietnam 2006-2008

2: Poverty Dynamics from Pseudo-Panel Data and Actual Data, Vietnam 2006- 2008

Poverty status 2006--> 2008	DLL M lower bound	Parametric		Truth	Parametric upper Model 1	DLLM upper bound	Parametric estimates		
		lower Model 1					$\rho= 0.5$	true ρ (0.6)	$\rho= 0.8$
Poor, Poor	12.1	12.5		7.6 (0.4)	4.8	6.1	7.1	7.7	9.2
Poor, Nonpoor	1.6	1.2		6.3 (0.4)	8.8	8.9	6.5	5.9	4.4
Nonpoor, Poor	0.2	0.2		4.3 (0.3)	7.9	8.1	5.6	5.0	3.4
Nonpoor, Nonpoor	86.1	86.1		81.9 (0.6)	78.5	76.8	80.8	81.4	82.9

Validation of Synthetic Panel Approach to Measuring Vulnerability in Vietnam

- Re-estimate (approach 2) vulnerability line
 - Original, true, vulnerability line = D5,480,000
 - Synthetic panel approach = D5,500,000

Validation: Vietnam

		2008				
		Poor	Vulnerable	Middle class	Total	
Panel A: True Panels	2006	Poor	9.9	5.4	0.4	15.7
			(0.8)	(0.5)	(0.2)	(0.9)
		Vulnerable	4.7	33.3	9.5	47.5
			(0.5)	(1.0)	(0.7)	(1.1)
		Middle class	0.2	9.4	27.1	36.8
			(0.1)	(0.7)	(0.9)	(1.0)
		Total	14.8	48.2	37.0	100
	(0.9)	(1.2)	(1.1)			
		2008				
		Poor	Vulnerable	Middle class	Total	
Panel B: Synthetic Panels	2006	Poor	10.8	5.3	0.1	16.2
			(0.3)	(0.1)	(0.0)	(0.3)
		Vulnerable	4.8	31.0	7.8	43.6
			(0.1)	(0.2)	(0.1)	(0.2)
		Middle class	0.1	9.3	30.8	40.2
			(0.0)	(0.1)	(0.3)	(0.4)
		Total	15.7	45.7	38.6	100.0
	(0.4)	(0.2)	(0.4)			

Trends: US, Vietnam, India

	US		Vietnam		India	
	2004	2008	2004	2008	2004	2009
Poor	8.9	10.0	20.0	14.3	36.8	31.5
Vulnerable	26.4	27.9	51.0	49.7	46.3	48.5
Middle class	64.7	62.1	29.0	36.0	16.8	20.0

Note: cut-off risk for India is set at 20%, not 10%

Vulnerability lines:

US = $2.77 * Z$

Vietnam = $2.14 * Z$

India = $2.05 * Z$

Conclusions

- We consider an alternative 3-way breakdown of the population: poor, non-poor but vulnerable, non-poor and secure
- We focus on the definition of a “vulnerability line” that would permit such a population breakdown
 - Try to link the line explicitly to a notion of risk of falling into poverty
- Propose that rising “prosperity” occurs when poverty *and* vulnerability decline

Conclusion

- The two alternative approaches we propose are simple to apply with panel data
 - Make “light” use of panel data
- Of course, panel data are rare, particularly in the developing world
- However, parallel research has pointed to ways to construct “synthetic panels” with cross-section surveys
 - Growing experience suggests method works quite well for a number of applications
- Synthetic panels can readily accommodate the vulnerability line calculations proposed here

Conclusion

- Empirical application to US, Vietnam and India in the period 2004-2008/9 point to some interesting findings:
 - “prosperity” in the US declined, increased in Vietnam
 - “vulnerability” in the US increased, declined in Vietnam.
 - poverty in US rose, fell in Vietnam
- Using a different “cut-off” level (20%) the picture in India points to:
 - falling poverty, rising vulnerability → “mixed” evolution of prosperity