## **Economic Mobility in Latin America**

Francisco H. G. Ferreira The World Bank and IZA

> University of Passau 3 December 2012

This presentation draws primarily on joint work (in progress) with Maria Ana Lugo. It also draws on Ferreira, Messina, Rigolini, López-Calva, Lugo and Vakis (2012): *Economic Mobility and the Rise of the Latin American Middle Class* (Washington, DC: World Bank)

### Outline

- 1. Concepts of economic mobility
- 2. Intergenerational mobility
  - Educational attainment and achievement
- 3. Intragenerational mobility
  - A 'matrix' decomposition
  - A detour: defining and measuring the middle class
  - Application to actual and synthetic panels
- 4. Concluding remarks

http://www.cnn.com/video/#/video/bestoftv/2012/11/26/exp-gps-1125-witw.cnn

- Sociological and economic approaches to mobility.
- Economic approach: studies the transformation of an initial income vector into one or more subsequent vectors, while keeping track of the identity of recipient units.
- Within the economic approach, three broad conceptions of mobility\*:



Figure 3: Mobility Concepts

\* Drawing on a taxonomy by Fields (2001)

These six concepts do capture very different aspects of 'mobility':

Examples:

	High	No
(1, 10, 100) $\longrightarrow$ (2, 20, 200)	IM (D & ND)	SM, PM, MTI, ELTI
(1, 10, 100) $\longrightarrow$ (100, 10, 1)	IM (ND), PM, SM, ELTI	IM (D)
(1, 10, 100) $\longrightarrow$ (36, 37, 38)	IM (ND), SM, ELTI	IM (D), PM

- These multiple ways of summarizing the information contained in the transition from one income vector to another when identities are preserved mirror the myriad measures of poverty and inequality.
  - Changes in poverty and inequality also summarize transitions from one vector to another, but with anonymity.
- Changes in some (anonymous) measures of poverty and inequality are ultimately simply different ways of aggregating the information contained in the *growth incidence curve*...

E.g. for the class of poverty measures that can be written as

$$P_t = \int_{-\infty}^{F(z)} \pi(y_t(p), z) dp$$

$$dP_t = \int_{-\infty}^{F(z)} \eta_t(p)g_t(p)dp + \pi(z,z)dF_t(z)$$

 For inequality measures – like the Gini coefficient, or the G.E. class – that can be written as:

$$I_t = G\left[\int_0^1 h\left(\frac{y_t(p)}{\mu}\right) dp\right]$$

- We have  $dI_t = G'(-) \int_0^1 h' \left(\frac{y_t(p)}{\mu}\right) \frac{\mu}{y_t(p)} \left[g_t(p) \frac{d\mu}{\mu}\right] dp$
- Can different (non-anonymous) measures of mobility also be expressed as aggregating information in some function, analogous to the GIC?
  - Answer: 'mobility profiles' (van Kerm, 2006, 2009).

 For inequality measures – like the Gini coefficient, or the G.E. class – that can be written as:

$$I_t = G\left[\int_0^1 h\left(\frac{y_t(p)}{\mu}\right) dp\right]$$

- We have  $dI_t = G'(-) \int_0^1 h' \left(\frac{y_t(p)}{\mu}\right) \frac{\mu}{y_t(p)} \left[g_t(p) \frac{d\mu}{\mu}\right] dp$
- Can different (non-anonymous) measures of mobility also be expressed as aggregating information in some function, analogous to the GIC?
  - Answer: 'mobility profiles' (van Kerm, 2006, 2009).

- For (all five sub-concepts within) the first two broad conceptions of mobility – movement and origin independence – the problem of measuring the overall extent of mobility in a society can be decomposed into two steps, in this order:\*
  - 1. Definition of an individual mobility function

$$m(p_0) = d(y_o(p_0), y_1(p_0))$$

2. Aggregation across individuals: Social mobility function

$$M(Y_0, Y_1) = \int_0^1 m(p_0) dp_0 = \int_0^1 d(y_0(p_0), y_1(p_0)) dp_0$$

\*See van Kerm (2006, 2009)



 Which is the integral of the na-GIC (Grimm, 2007) along initial ranks.

\* The log-approximation of this measure is the M<sup>3</sup> measure in Fields and Ok (1999) and Fields et al. (2002).

Concept	m(p <sub>0</sub> ) - example	Profile (Peru, 2004-2006)
Directional income movement	$d(y_0, y_1) = y_1 - y_0$ $d(y_0, y_1) = \frac{y_1 - y_0}{y_0}$	Differences in log-income
Non-directional income movement	$d(y_0, y_1) =  y_1 - y_0 $ $d(y_0, y_1) = \left \frac{y_1 - y_0}{y_0}\right  \approx  \log y_1 - \log y_0 $	Absolute proportional differences in income
Share movement	$d(y_0, y_1) = \left  \frac{y_1}{\mu_1} - \frac{y_0}{\mu_0} \right $	Absolute difference in income shares

Concept	m(p <sub>0</sub> ) - example	Profile (Peru, 2004-2006)
Positional movement	$d(y_0, y_1) =  rank_1 - rank_0 $	State of the second sec
Mobility as origin (or time) independence	$d(y_0, y_1) = \frac{1}{2} \left( \frac{y_0 - \mu_0}{s_0} - \frac{y_1 - \mu_1}{s_1} \right)^2$ $M(Y_0, Y_1) = 1 - \rho_{y_0 y_1}$ (D'Agostino and Dardanoni, 2006)	Square differences in standardized income
Mobility as an equalizer of long-term incomes	Is not an aggregate of individual mobility functions.	$1 - \frac{I\left(\underline{y_0} + \underline{y_1}\right)}{I\left(\underline{y_0}\right)}$

### • Pick one per domain:

Concept \ Domain	Intra-generational	Inter-generational
Directional Income Movement	<ul> <li>Growth in individual or household incomes , and well-being.</li> <li>Movements in and out of poverty, and the middle class.</li> </ul>	- 'Abolute' progress between generations: how much better off are the children than the parents?
Mobility as origin (or time) independence	- Long-term life-cycle movements: life time achievements independent of initial conditions.	-Equal opportunities: Children's achievements independent of parent's circumstances.

• **Empirical Challenge**: Scarcity of panel data for intra-generational mobility, and of systematic information on the parents of today's adults.

- The concept of interest is mobility as origin independence.
- Take the mobility profile given by (half) the square of the difference between standardized incomes:

$$d(y_0, y_1) = \frac{1}{2} \left( \frac{y_0 - \mu_0}{s_0} - \frac{y_1 - \mu_1}{s_1} \right)^2$$

- The corresponding social mobility function is  $M(Y_0, Y_1) = 1 \rho_{y_0y_1}$
- Which is the complement to the square root of the R<sup>2</sup> in the old Galtonian regression:

$$y_t = \beta y_{t-1} + \varepsilon_t$$

$$R^{2} = \frac{Var(\beta y_{t-1})}{Var(y_{t})} = \frac{Cov^{2}(y_{t-1}, y_{t})}{Var(y_{t-1})} = \rho_{t,t-1}^{2}$$

- A distinguished international literature has examined these correlations in long-term panels with information on earnings for parents and (typically) sons:
  - Björklund and Jäntti (AER, 1997) : US and Sweden
  - Couch and Dunn (JHR, 1997): US and Germany
  - Dearden, Machin and Reed (EJ, 1997): UK
  - Solon (*JEP*, 2002): cross country
  - Mazumder (*REStat*, 2005): US
- In LAC, studies have either relied on education, or on TSIV:
  - Behrman, Gaviria and Székely (*Economía*, 2001)
  - Gaviria (*Economía*, 2007)
  - S. Ferreira and Veloso (PPE, 2003; BRE 2006)

#### COUNTRIES RANKED BY AVERAGE PARENT-CHILD SCHOOLING CORRELATION, AGES 20-69

## 2. Intergenerational mobility

The correlation coefficient has also been used to measure mobility in educational **attainment** 

The correlation between years of schooling of parents and children

Source: Hertz et al. (2007)

Country	Coefficient	Rank	Correlation	Rank
Peru	0.88	6	0.66	1
Ecuador	0.72	12	0.61	2
Panama	0.73	11	0.61	3
Chile	0.64	18	0.60	4
Brazil	0.95	4	0.59	5
Colombia	0.80	8	0.59	6
Nicaragua	0.82	7	0.55	7
Indonesia	0.78	9	0.55	8
Italy†	0.67	17	0.54	9
Slovenia†	0.54	27	0.52	10
Egypt	1.03	2	0.50	11
Hungary†	0.61	20	0.49	12
Sri Lanka	0.61	19	0.48	13
Pakistan	1.00	3	0.46	14
USA	0.46	33	0.46	15
Switzerland†	0.49	30	0.46	16
Ireland†	0.70	15	0.46	17
South Africa (KwaZulu-Natal)	0.69	16	0.44	18
Poland†	0.48	31	0.43	19
Vietnam	0.58	23	0.40	20
Philippines	0.41	36	0.40	21
Belgium (Flanders)	0.41	35	0.40	22
Estonia	0.54	28	0.40	23
Sweden	0.58	26	0.40	24
Ghana	0.71	13	0.39	25
Ukraine	0.37	40	0.39	26
East Timor	1.27	1	0.39	27
Bangladesh (Matlab)	0.58	25	0.38	28
Slovakia	0.61	21	0.37	29
Czech Republic†	0.44	34	0.37	30
The Netherlands	0.58	24	0.36	31
Norway	0.40	38	0.35	32
Nepal	0.94	5	0.35	33
New Zealand†	0.40	37	0.33	34
Finland	0.48	32	0.33	35

And if we look at achievements, things do not look much better



Relationship of average PISA test scores and intergenerational mobility across 65 countries and economies, 2009

### Differences in the educational gap between the top and bottom income quintiles in Latin America, 1995–2009

There have been slight improvements over the last two decades...



...but family background remains too important (and some of that effect operates through school sorting)

#### Direct and overall impact of parental background on children's test scores



- Despite the paucity of longer panels, there is a large(ish) literature on mobility in LAC. Examples:
  - Actual panels:
    - Scott (2000)
    - Beccaria and Groisman (2006)
    - Contreras et al. (2006)
    - Fields, Duval, Freije and Sanchez-Puerta (2007)
    - Grimm (2007)
  - Pseudo-panels
    - Antman and McKenzie (2007)
    - Calonico (2006)

- Now the concept of interest (for us) is directional income movement.
- A generalization of  $M^3(Y_0, Y_1)$ , which gives the integral of the na-GIC, is a suitable measure of (proportional) directional income movement.

$$M_{\alpha}(Y_{0},Y_{1}) = \int_{0}^{1} \left(\frac{y_{1}(p_{0}) - y_{0}(p_{0})}{y_{0}(p_{0})}\right)^{\alpha} dp_{0} = \int_{0}^{1} g(p_{0})^{\alpha} dp_{0} \quad \alpha \in (0,1)$$
Proportional differences in income
  
• This index can be decomposed
'horizontally', into gainers and losers:
$$m(Y_{0},Y_{1};\alpha) = \int_{0}^{1} m(y_{0}(p_{0}),y_{1}(p_{0});\alpha) dp$$

$$\int_{d(y_{0},y_{1})<0} m(y_{0}(p_{0}),y_{1}(p_{0});\alpha) dp + \int_{d(y_{0},y_{1})>0} m(y_{0}(p_{0}),y_{1}(p_{0});\alpha) dp$$
Figure 3. Horizontal decomposition of the mobility profile (proportional income changes). Peru 2004-06.

.2

.8

.6

Percentile

0

Or 'vertically', by 'class' or origin:

$$M(Y_0, Y_1; \alpha) = \int_{0}^{F(z)} m(y_0, y_1; \alpha) dp + \int_{F(z)}^{F(\zeta_L)} m(y_0, y_1; \alpha) dp + \int_{F(\zeta_L)}^{F(\zeta_H)} m(y_0, y_1; \alpha) dp + \int_{0F(\zeta_H)}^{F(\zeta_H)} m(y_0, y_1; \alpha) dp$$

$$= q_P \overline{m_p}(y_0, y_1; \alpha) + q_V \overline{m_V}(y_0, y_1; \alpha) + q_M \overline{m_{MC}}(y_0, y_1; \alpha) + q_R \overline{m_R}(y_0, y_1; \alpha),$$

Figure 4. Vertical decomposition of the mobility profile (proportional income changes). Peru 2004-06.



- An application: partition the distribution into 'economic classes', by analogy to the identification procedure of Sen (1976) for unidimensional poverty:
  - Poverty: a state where the basic functionings of <u>food</u> security and good nutrition are not guaranteed.
    - z = PPP\$4/day per capita
  - Middle-class: a state defined by the basic functioning of <u>economic security</u> (proxied by low vulnerability to falling back into poverty)
    - "Validation" using a complementary approach: <u>subjective self-assessment.</u>
    - $\zeta_L = PPP$10/day per capita$
  - Elite: the politically powerful top of the distribution
    - $\zeta_{\rm H} = PPP$50/day per capita$

Vulnerability to poverty: five year intervals in a true panel



Source: López-Calva and Ortiz-Juarez (2011)

- Opinion and value surveys (*Ecosocial*) contain information on selfreported social class, and on household assets (but no income information)
- Using an approach similar to poverty mapping, we impute permanent incomes into *Ecosocial* 
  - Using the coefficients from an income regression on household assets, estimated using ancillary household surveys
- Finally, we plot the non-parametric density functions of self-reported social class against predicted permanent income, and define the lower middle class threshold as the crossing point between those who see themselves as lower middle class or poor, and those who see themselves as "middle-middle" or above.

Self-reported social classes: Mexico (2007)



Country	Lower MC threshold 2005 USD PPP	Income percentile
Brazil	16.3	84
Chile	20.3	83
Colombia	9.25	69
Mexico	9.6	68
Peru	10.5	76

- Thresholds range from US\$ 9.3 to 20.3 per day
  - This confirms that middle-class perceptions are country-specific.
- But the US\$ 10 line obtained from the vulnerability approach appears to be consistent with a lower envelope for these 'subjective thresholds'.
  - Note that this threshold is relatively high in the income distribution.



Four economic classes, by income distribution, in selected Latin American countries

Average per capita income in the US, 2009

• The two decompositions are additive and can be combined into a 'matrix decomposition':

$$\begin{split} M(Y_0,Y_1;\alpha) &= q_P^p \overline{m_P^p}(y_0,y_1;\alpha) + q_P^v \overline{m_P^v}(y_0,y_1;\alpha) + q_P^{MC} \overline{m_P^{MC}}(y_0,y_1;\alpha) + q_P^R \overline{m_P^R}(y_0,y_1;\alpha) + \\ q_V^v \overline{m_V^p}(y_0,y_1;\alpha) + q_V^v \overline{m_P^v}(y_0,y_1;\alpha) + q_V^{MC} \overline{m_W^{MC}}(y_0,y_1;\alpha) + \\ q_M^P \overline{m_{MC}^P}(y_0,y_1;\alpha) + q_M^v \overline{m_{MC}^v}(y_0,y_1;\alpha) + q_{MC}^{MC} \overline{m_{MC}^{MC}}(y_0,y_1;\alpha) + \\ q_R^P \overline{m_R^P}(y_0,y_1;\alpha) + q_R^v \overline{m_N^v}(y_0,y_1;\alpha) + q_R^{MC} \overline{m_{MC}^{MC}}(y_0,y_1;\alpha) + \\ q_R^P \overline{m_R^P}(y_0,y_1;\alpha) + q_R^v \overline{m_R^v}(y_0,y_1;\alpha) + q_R^{MC} \overline{m_R^{MC}}(y_0,y_1;\alpha) + \\ q_R^P \overline{m_R^P}(y_0,y_1;\alpha) + q_R^v \overline{m_R^v}(y_0,y_1;\alpha) + q_R^{MC} \overline{m_R^{MC}}(y_0,y_1;\alpha) + \\ q_R^P \overline{m_R^P}(y_0,y_1;\alpha) + q_R^v \overline{m_R^v}(y_0,y_1;\alpha) + q_R^{MC} \overline{m_R^{MC}}(y_0,y_1;\alpha) + \\ q_R^P \overline{m_R^P}(y_0,y_1;\alpha) + q_R^v \overline{m_R^v}(y_0,y_1;\alpha) + \\ q_R^{MC} \overline{m_R^{MC}}(y_0,y_1;\alpha) + q_R^W \overline{m_R^P}(y_0,y_1;\alpha) + \\ q_R^W \overline{m_R^P}(y_0,y_1;\alpha) + q_R^V \overline{m_R^v}(y_0,y_1;\alpha) + \\ q_R^W \overline{m_R^P}(y_0,y_1;\alpha) + q_R^W \overline{m_R^v}(y_0,y_1;\alpha) + \\ q_R^W \overline{m_R^P}(y_0,y_1;\alpha) + q_R^W \overline{m_R^V}(y_0,y_1;\alpha) + \\ q_R^W \overline{m_R^P}(y_0,y_1;\alpha) + q_R^W \overline{m_R^V}(y_0,y_1;\alpha) + \\ q_R^W \overline{m_R^V}(y_0,y_1;\alpha)$$

 When α = 0, this decomposition is the sum of all cells in a transition matrix (with cell boundaries given by fixed income thresholds):

$$M(Y_0, Y_1; 0) = q_P^P + q_P^V + q_P^{MC} + q_P^R + q_V^{P} + q_V^{V} + q_V^{MC} + q_V^R + q_M^{P} + q_M^{P} + q_M^{P} + q_M^{V} + q_M^{MC} + q_M^R + q_M^{P} + q_R^{V} + q_M^{R} + q_M^{R} + q_R^{R} + q_R^$$

There are a number of interesting cuts at this decomposition. This one focuses on movers and stayers...

Origin\Destination	Poor	Near Poor	MC & above
Poor	Chronic Poverty	U	o <sub>wardly mobile</sub>
Near Poor	Dou	Near Poor	
MC & above	-own <sub>wardly</sub> mob	ile	Established Middle Class

This one focuses on poverty transitions:

Origin\Destination	Poor	Near Poor	MC & above
Poor	Chronic Poverty	Poverty	leavers
Near Poor	Entrants	Near Poor (Vulnerable)	
MC & above	poverty		

This one focuses on middle-class transitions:

Origin\Destination	Poor	Near Poor	MC & above
Poor			
			The New
Near Poor			Middle Class
MC & above	The Dis	splaced	Established Middle Class

• An application to Peru: 2004-2006 ( $\alpha = 0$ )



2,006 (Destination)

\* Numbers in the cells are percentages.

• An application to Peru: 2004-2006

### Proportional income changes in each cell

		Р	V	MC+	
	Ρ	.28	1.34	4.21	.77
2,004	V	44	.13	1.37	.36
(origin)	MC+	79	47	.21	02
		.10	.32	.80	.40
		.10	.32	.80	.40

#### 2,006 (Destination)

• An application to Peru: 2004-2006 ( $\alpha = 1$ )



2,006 (Destination)

The decomposition of  $M(Y_{o}, Y_{1}, \alpha=1)$  yields the product of the previous two matrices: population proportions \* mean income growth per cell.

- Lanjouw, Luoto and McKenzie (2011): a new approach to using information from repeated cross-sections to estimate bounds on aggregate economic mobility:
  - 1. Using two cross-sectional surveys, for years  $t_1\,\text{and}\,t_2$
  - 2. We observe income  $y_{i1}$  for households in year  $t_1$ . How can we predict income for these households in year  $t_2$ ?
  - 3. First, estimate the relationship between incomes and **time-invariant** correlates in each year  $y_{i1} = \beta_1 ' x_{i1} + \varepsilon_{i1}$   $y_{i2} = \beta_2 ' x_{i2} + \varepsilon_{i2}$
  - 5. Then, predict incomes for  $t_1$  households in year  $t_2$  using the same set of timeinvariant characteristics **and** the returns estimated at  $t_2$ . Different assumptions about the residuals yield lower and upper income bounds.
  - Lower bound (perfect correlation)
  - Upper bound (no correlation)

$$\hat{y}_{i2}^{1I} = \hat{\beta}_2 \, \mathbf{x}_{i1} + \hat{\varepsilon}_{i1}$$

$$\hat{y}_{i2}^{1S} = \hat{\beta}_2 \, \mathbf{x}_{i1} + \hat{\varepsilon}_{i2}$$

- These lower and upper bounds on individual mobility are generally imprecise.
- But when aggregated across (sub-) populations, they **can** yield meaningful intervals.
- Validation exercises for actual panels are generally supportive.
- Because of measurement error, lower bound estimates are not necessarily worse then panel estimates.
- Lower bound estimates can generate (transition) matrix decompositions analogous to those from Peru's actual panel above.
- What follows are somewhat coarser examples, with estimates of transitions into and out of poverty and the middle class.

## LAC experienced high levels of upward mobility in the past 15 years...

## Intragenerational mobility in Latin America over the past 15 years (circa 1995–2010): lower bounds

#### percentage of population

			Destination (c. 2010)			
		Poor	Vulnerab	le Middle cla	ss Total	
	Poor	22.5	21.0	2.2	45.7	
Origin (c.1995)	Vulnerable	0.9	14.3	18.2	33.4	
	Middle class	0.1	0.5	20.3	20.9	
Total		23.4	35.9	40.7	100.0	

...although there is considerable heterogeneity.

#### Intragenerational mobility in Latin America, by country



## Education is a strong predictor of ascending into the middle class

Upward mobility conditional on initial education, by country



The cumulative result of these mobility patterns over the last fifteen years has been a remarkable social transformation

Middle class, vulnerability, and poverty trends in Latin America, 1995–2009



### 4. Concluding remarks

- At least six different meanings for economic mobility
  - Five of those can be underpinned by mobility profiles built upon individual mobility functions
- We choose to focus on:
  - mobility as origin independence to study intergenerational mobility
  - Mobility as directional income movement to study intragenerational mobility
- Evidence that mobility across generations is low in LAC exists for educational attainment and achievement
  - There is some evidence of a recent improvement.
  - But less so for achievement than for attainment.

### 4. Concluding remarks

- A standard measure of mobility as (proportional) directional income movement is simply the integral of the non-anonymous growth incidence curve (na-GIC).
- This measure can be 'matrix-decomposed' into terms corresponding to the upward and downward movements into and out of specific "economic classes", such as the poor, the vulnerable or the middle class.
- Estimates of this decomposition, both for actual and "pseudo" panels, suggest that there has been considerable movement out of poverty and into the middle class in Latin America in the last decade or two
  - There has continued to be some offsetting downward movement too
  - Education and labor market formality are correlated with those movements.