Inequality of Opportunity and Economic Growth: Can cross-country regressions tell us anything?

Francisco H. G. Ferreira*#
Christoph Lakner*
Maria Ana Lugo*
Berk Özler*

* The World Bank # IZA,Bonn

World Statistics Congress
Rio de Janeiro, 29 July 2015

Outline

1. Motivation
   - Inequality and growth: a review of the theory
   - Inequality and growth: a review of the evidence
2. Inequality of opportunity
3. Data
4. Econometric specification
5. Results
   - Income and expenditure surveys
   - Demographic and Health Surveys
6. Conclusions

Inequality and growth: a review of the theory

- Inequality has been hypothesized to affect economic growth through various mechanisms:
  - Savings
    - Kaldor (1957)
  - Credit constraints and investment indivisibilities
    - Banerjee and Newman (1993)
    - Galor and Zeira (1993)
    - Aghion and Bolton (1997)
  - Political economy
    - Alesina and Rodrik (1994)
    - Persson and Tabellini (1994)
    - Bénabou (2000)
    - Ferreira (2001)
  - Excellent surveys include:
    - Voitchovsky (2009)

Inequality and growth: a review of the evidence

- Phase I - cross-section results: "inequality is bad for growth"

- Alesina and Rodrik (1994)
- Persson and Tabelini (1994)
- Deininger and Squire (*JDE*, 1998)

Table from Alesina and Rodrik (1994) →

TABLE I
GROWTH REGRESSIONS FOR 1960–1985

	High-quality sample ($N = 46$)		Largest possible sample ($N = 70$)		Largest possible sample ($N = 49$)		Largest possible sample ($N = 41$)	
	OLS (1)	TSLS (2)	OLS (3)	TSLS (4)	OLS (5)	OLS (6)	OLS (7)	OLS (8)
Const.	3.60	8.66	1.76	6.48	3.71	6.22	6.24	6.21
	(2.66)	(3.33)	(1.50)	(2.93)	(3.86)	(4.69)	(4.63)	(4.61)
GDP60	−0.44	−0.52	−0.48	−0.58	−0.38	−0.38	−0.39	−0.38
	(−3.28)	(−3.17)	(−3.37)	(−3.47)	(−3.61)	(−3.25)	(−3.06)	(−2.95)
PRIM60	3.26	2.85	3.98	3.70	3.85	2.66	2.62	2.65
	(3.38)	(2.43)	(4.66)	(3.72)	(4.88)	(2.66)	(2.53)	(2.56)
GINI60	−5.70	−15.98	3.58	−12.93		−3.47	−3.45	−3.47
	(−2.46)	(−3.21)	(−1.81)	(−3.12)		(−1.82)	(−1.79)	(−1.80)
GINILND					−5.50	−5.23	−5.24	−5.21
					(−5.24)	(−4.38)	(−4.32)	(−4.19)
DEMOC* GINILND							0.12	
							(0.12)	
DEMOC								0.02
								(0.05)
$\bar{R}^2$	0.28	0.27	0.25	0.26	0.53	0.53	0.51	0.51

The dependent variable is average per capita growth rate over 1960–1985. *t*-statistics are in parentheses. Independent variables are defined as follows:

GDP60: Per capita GDP level in 1960
PRIM60: Primary school enrollment ratio in 1960
GINI60: Gini coefficient of income inequality, measured close to 1960 (see Appendix for dates)
GINILND: Gini coefficient of land distribution inequality, measured close to 1960 (see Appendix for dates)
DEMOC: Democracy dummy.

Two-stage least squares regressions use GDP60, PRIM60, literacy rate in 1960, infant mortality in 1965, secondary enrollment in 1960, fertility in 1965, and an Africa dummy as instruments.

Inequality and growth: a review of the evidence

- Forbes (AER, 2000): With panel data, (recent) inequality is good for growth:

TABLE 3—REGRESSION RESULTS: ALTERNATE ESTIMATION TECHNIQUES

	Five-year periods				Ten-year periods:
Estimation method	Fixed effects (1)	Random effects (2)	Chamberlain's π-matrix (3)	Arellano and Bond (4)	fixed effects (5)
Inequality	0.0036	0.0013	0.0016	0.0013	0.0013
	(0.0015)	(0.0006)	(0.0002)	(0.0006)	(0.0011)
Income	−0.076	0.017	−0.027	−0.047	−0.071
	(0.020)	(0.006)	(0.004)	(0.008)	(0.016)
Male Education	−0.014	0.047	0.018	−0.008	−0.002
	(0.031)	(0.015)	(0.010)	(0.022)	(0.028)
Female Education	0.070	−0.038	0.054	0.074	0.031
	(0.032)	(0.016)	(0.006)	(0.018)	(0.030)
PPP	−0.0008	−0.0009	−0.0013	−0.0013	−0.0003
	(0.0003)	(0.0002)	(0.0000)	(0.0001)	(0.0003)
R^2	0.67	0.49			0.71
Countries	45	45	45	45	45
Observations	180	180	135	135	112
Period	1965–1995[a]	1965–1995[a]	1970–1995	1970–1995	1965–1995

Notes: Dependent variable is average annual per capita growth. Standard errors are in parentheses. R^2 is the within-R^2 for fixed effects and the overall-R^2 for random effects.

[a] Estimates are virtually identical for the period 1970–1995 (with 135 observations).

Inequality and growth: a review of the evidence

- Phase II:
  - Li and Zou (RDE, 1998):
    - Similar results to Forbes.
  - Barro (*JEG*, 2000):
    - No overall relationship, but negative effect of inequality on growth among poor countries, and positive among rich countries.
  - Voitchovsky (*JEG*, 2005):
    - No overall relationship when inequality is measured by a scalar index, but top-end inequality "good", while bottom-end inequality "bad".
  - Banerjee and Duflo (*JEG*, 2003): "The …data does seem inconsistent with a linear structure."
- Phase III:
  - Easterly (2007):
    - Inequality, instrumented by agricultural endowments, hurts growth
  - Berg, Ostry and Zettelmeyer (JDE, 2012):
    - Inequality reduces the duration of high-growth spells
  - Ravallion (AER, 2012):
    - Initial poverty, rather than inequality, is negatively associated with economic growth (and also with the growth elasticity of poverty)

Are we capturing the "right" inequality?

- Lack of subtlety in the basic cross-section approach often criticized: top vs. bottom, poor vs. rich countries; inequality vs. poverty, etc.
- A new angle: the cholesterol analogy
  - Inequality of opportunity vs. inequality of effort

"The rise in inequality in the United States over the last three decades has reached the point that inequality in incomes is causing an unhealthy division in opportunities, and is a threat to our economic growth."

(Alan Krueger, 12 January 2012)

- How can we empirically separate out inequality of opportunity from that due to efforts?

Inequality of Opportunity

- Large literature on definitions and measurement of inequality of opportunity
  - Roemer (1993, 1998)
  - Van de Gaer (1993)
  - Bourguignon et al. (2007)
  - Checchi and Peragine (2010)
  - Ferreira and Gignoux (2011)
  - Recent surveys: Roemer and Trannoy (2015) and Ferreira and Peragine (forthcoming)

- Basic framework:

Population of agents indexed by i, $i \in \{1,...N\}$, with i characterized by $\{y_i, C_i, e_i\}$

C is a vector of *J* elements, each of which is discrete and can take a finite number of values x_j.

Define a partition $\Pi = \{T_1, T_2, ..., T_K\}$

such that $C_i = C_j, \forall i, j | i \in T_k, j \in T_k, \forall k.$

Inequality of Opportunity

- Following Van de Gaer (1993), Ooghe et al. (2007) and Ferreira and Gignoux (2011), we consider the support of $F^k(y)$ as each type's opportunity set.
- Value it by its mean: $\mu^k(y) = \int_0^\infty y dF^k(y)$
- Then equality of opportunity is attained when $\mu^k(y) = \mu^l(y), \forall l,k | T_k \in \Pi, T_l \in \Pi$
- Measuring inequality of opportunity would now seem to require assessing differences in the smoothed distribution $\{\mu_i^k\}$, rather than on the marginal distribution y = $(y_1,\ldots,y_N)$
- This immediately leads to two candidate measures, which yield lower-bound measures:
- IOL: $\theta_a = I(\{\mu_i^k\})$ and IOR: $\theta_r = \frac{I(\{\mu_i^k\})}{I(y)}$
- If *I()* is one of the Theil indices, then θ_a is the between-group component of a standard Theil decomposition, and θ_r is R_B (Cowell and Jenkins, 1995).

Inequality of Opportunity - Precedent

- Marrero and Rodriguez (2013): "Inequality of opportunity and growth", *Journal of Development Economics*, **104**: 107-122.
  - Use PSID data for 26 US states
  - t=3: 1970, 1980, 1990
  - Ex-ante I.Op.: between-type inequality
  - Eight types, on the basis of race and father's education
  - Household heads aged 18-65 only.

Table 4. Inequality, IO and growth within US States: 1980-2000

	(1) for Theil 0 and IO (8 groups)	(2) for Theil 0 and IO (8 groups)	(1) for Gini and IO (8 groups)	(2) for Gini and IO (8 groups)
Lagged per capita income	-0.0018*** (0.00004)	-0.0017*** (0.00012)	-0.0020*** (0.00006)	-0.0018*** (0.00023)
Inequality	-2.9524 (4.7388)	10.7953*** (0.4273)	13.7357* (7.1776)	14.8684*** (5.1708)
Inequality of opportunity	-12.5087*** (4.4638)	-26.3222*** (1.3883)	-17.3267*** (0.14066)	-9.3317*** (0.6125)
Dum 80	0.4016 (0.3760)	0.9122 (0.5877)	0.0018 (0.4092)	0.5339 (0.3685)

Data

- In this paper, we explore the relationship between IOp and economic growth in two panels of countries that are as close to that used by Forbes (2000) as possible:
  - Income / expenditure surveys (IES)
    - Advantages: nationally representative, detailed income / consumption indicators
    - Disadvantages: only 19 developing countries (out of 42)
  - Demographic and Health Surveys (DHS)
    - Advantages: 42 developing countries, broader geographical coverage
    - Disadvantages: measure of economic advantage is a wealth index.

Data: Income / expenditure surveys

- Unlike Forbes (2000), we had to construct IOp measures from household-level data.
- GMM specification requires comparable surveys across at least three time periods.
  - 42 countries, 118 surveys
- Sources: Luxembourg Income Study, World Bank (I2D2), SEDLAC, three additional surveys.
- Partition into types is held constant within countries over time
- Wide diversity of data availability on circumstances across countries
  - Number of types range from 6 to 1224 (Table A1)
  - Dummy included for income v. expenditure
  - Quartic on number of types included.

Data: Demographic and Health Surveys

- Again: GMM specification requires comparable surveys across at least three time periods.
  - 42 countries, 134 surveys
- IOp measure: $\theta_{DHS} = Var(\{C_i\hat{\psi}\})$ obtained from a regression $y_i = C_i\psi + \varepsilon_i$
- where $y_i = \sum_{f=1}^{F} a_f \left(\frac{x_{fi} - \bar{x}_f}{s_f} \right)$ is the first principal component of a set of household assets, *x*.
  - Ferreira, Gignoux and Aran (JOEI, 2011)
  - Filmer and Pritchett (Dem, 2001)
  - McKenzie (J. Pop. E, 2005)
- Circumstances include religion, ethnicity, mother tongue, region of birth, number of siblings.

Econometric Specification

- On both the IES and DHS samples we estimate:

$$g_{it} = \beta_1 y_{i,t-5} + \beta_2 I(y)_{i,t-5} + \beta_3 ME_{i,t-5} + \beta_4 FE_{i,t-5} + \beta_5 PPPI_{i,t-5} + \alpha_i + \eta_t + u_{it}$$

- and

$$g_{it} = \beta_1 y_{i,t-5} + \beta_2 IOp_{i,t-5} + \beta_3 IR_{i,t-5} + \beta_4 ME_{i,t-5} + \beta_5 FE_{i,t-5} + \beta_6 PPPI_{i,t-5} + \alpha_i + \eta_t + u_{it}$$

  - IOp: between-type share of the MLD (for IES)
  - Controls:
    - Male education: Proportion of men 25 and older with any secondary schooling
    - Female education: Proportion of women 25 and older with any secondary schooling
    - PPPI: market distortions in the price level of investment goods relative to the US (Summers and Heston database, as used by Forbes, 2000)

Econometric Specification

- Both equations are estimated under the following specifications:
- OLS, fixed effects, long-run OLS (10 years), difference-GMM, system-GMM
  - OLS: stable workhorse, but regressors likely endogenous.
  - Fixed effects: Control for time-invariant unobservables, but lagged dependent variable makes FE estimator unreliable.
  - Difference GMM: Persistent time series + small T -> weak instrument biases.
  - System GMM: Instrument proliferation -> weak Hansen test (overidentification); use PCA and collapsing to reduce instrument set (Roodman, 2009).
  - System GMM with collapsed instrument set: use Kleinbergen-Papp test for under-identification (Bazzi and Clemens, 2013)

Results: Total inequality

Table 3. Economic growth on total inequality

Income/expenditure survey sample

						System-GMM		
	OLS	*FE*	*Long-run-OLS*	*Difference GMM*	*Full*	*Collapse*		
						System	*Difference equation*	*Levels equation*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Log initial GDP per capita	-0.005	-0.206***	-0.007	-0.153**	0.001	-0.028	-0.188***	0.002
	(0.004)	(0.051)	(0.006)	(0.060)	(0.022)	(0.046)	(0.058)	(0.104)
Total inequality (set 2) (lagged)	-0.037*	-0.174*	0.000	-0.129	-0.056	-0.303	-0.256	-0.092
	(0.021)	(0.092)	(0.020)	(0.168)	(0.064)	(0.249)	(0.193)	(0.421)
Female secondary education (lagged)	0.052	1.138**	-0.005	1.937	0.032	0.408	2.012**	0.563
	(0.049)	(0.516)	(0.060)	(1.503)	(0.127)	(0.478)	(0.898)	(0.719)
Male secondary education (lagged)	-0.021	-0.950	0.071	-1.462	0.013	-0.619	-1.569	-0.957
	(0.056)	(0.579)	(0.068)	(1.608)	(0.151)	(0.688)	(0.988)	(0.939)
Price level of investment (lagged)	-0.001***	0.000	-0.000	-0.001	-0.001*	-0.001	0.000	-0.001
	(0.000)	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.009)
Indicator of income data	-0.015	0.000	-0.020		-0.019	0.055		
	(0.010)	(.)	(0.015)		(0.044)	(0.100)		
Constant	0.156***	1.804***	0.112**		0.137	0.498		
	(0.037)	(0.429)	(0.043)		(0.139)	(0.435)		

Table 4. Economic growth on total inequality

Demographic and Health Survey sample

	OLS	*FE*	*Long-run-OLS*	*Difference GMM*	*System-GMM*			
					Full	*Collapse*		
						System	*Difference equation*	*Levels equation*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Log initial GDP per capita	-0.001	-0.138***	-0.006	-0.172***	-0.007	-0.040	-0.321***	-0.021
	(0.006)	(0.026)	(0.009)	(0.047)	(0.017)	(0.028)	(0.047)	(0.018)
Total inequality (lagged)	-0.001	0.016	-0.006	0.044	0.010	-0.034	0.019	-0.020
	(0.004)	(0.022)	(0.005)	(0.045)	(0.025)	(0.052)	(0.041)	(0.056)
Female secondary education (lagged)	0.053	0.284	-0.178	0.523	0.073	0.433	1.629	0.349
	(0.104)	(0.523)	(0.145)	(0.861)	(0.207)	(0.286)	(1.254)	(0.509)
Male secondary education (lagged)	-0.003	-0.236	0.217*	-0.048	0.018	-0.367	-1.549	-0.306
	(0.083)	(0.468)	(0.118)	(0.911)	(0.160)	(0.368)	(1.404)	(0.601)
Price level of investment (lagged)	-0.000	0.000	0.000	0.000	0.000	-0.000	0.000	-0.001
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.001)
Constant	0.011	0.790***	0.085		0.004	0.452		
	(0.057)	(0.202)	(0.061)		(0.200)	(0.425)		

Results: Inequality of opportunity

Table 5. Economic growth on inequality of opportunity and residual inequality

Income/expenditure survey sample

					System-GMM			
						Collapse		
	OLS	*FE*	*Long-run-OLS*	*Difference GMM*	*Full*	*System*	*Difference equation*	*Levels equation*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Log initial GDP per capita	-0.003	-0.224***	-0.007	-0.189***	-0.017	-0.023	-0.173***	0.031
	(0.005)	(0.050)	(0.006)	(0.061)	(0.015)	(0.025)	(0.052)	(0.054)
Inequality of Opp. (set 2) (lagged)	-0.070	-0.050	-0.072	-0.144	-0.234	-0.681	-0.400	-0.227
	(0.074)	(0.193)	(0.087)	(0.286)	(0.209)	(0.447)	(0.351)	(2.107)
Residual inequality (set 2) (lagged)	-0.036	-0.210	0.029	0.063	-0.063	-0.175	-0.143	0.054
	(0.035)	(0.145)	(0.075)	(0.353)	(0.095)	(0.240)	(0.251)	(0.856)
Female secondary education (lagged)	0.069	0.991*	-0.008	2.441	0.064	0.255	1.657**	0.588
	(0.046)	(0.497)	(0.081)	(1.501)	(0.110)	(0.424)	(0.805)	(0.667)
Male secondary education (lagged)	-0.052	-0.819	0.080	-2.026	-0.018	-0.506	-1.163	-0.988
	(0.055)	(0.563)	(0.111)	(1.558)	(0.137)	(0.670)	(0.850)	(1.020)
Price level of investment (lagged)	-0.001**	0.000	-0.000	-0.001	-0.000	-0.000	0.000	-0.003
	(0.000)	(0.001)	(0.000)	(0.001)	(0.000)	(0.001)	(0.001)	(0.005)
Indicator of income data	-0.023*	0.000	-0.026		-0.004	0.011		
	(0.011)	(.)	(0.016)		(0.039)	(0.085)		
Constant	0.143***	1.933***	0.102*		0.243**	0.460		
	(0.041)	(0.445)	(0.058)		(0.108)	(0.293)		

Table 6. Economic growth on inequality of opportunity and residual inequality

Demographic and Health Survey sample

					System-GMM			
	OLS	*FE*	*Long-run-OLS*	*Difference GMM*	*Full*	*Collapse*		
						System	*Difference equation*	*Levels equatio*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Log initial GDP per capita	-0.003	-0.137***	-0.005	-0.160***	-0.008	-0.021	-0.295***	-0.113
	(0.007)	(0.028)	(0.010)	(0.046)	(0.014)	(0.024)	(0.048)	(0.978)
Inequality of Opp. (lagged)	0.006	0.005	-0.017**	0.051	0.033	0.012	-0.018	-0.013
	(0.007)	(0.040)	(0.008)	(0.060)	(0.027)	(0.051)	(0.081)	(0.296)
Residual inequality (lagged)	-0.001	0.022	-0.001	0.028	0.014	-0.001	0.052	-0.397
	(0.006)	(0.031)	(0.007)	(0.024)	(0.028)	(0.043)	(0.054)	(3.859)
Female secondary education (lagged)	0.047	0.365	-0.176	0.550	-0.086	0.181	1.276	-1.344
	(0.106)	(0.621)	(0.150)	(1.155)	(0.256)	(0.286)	(1.206)	(15.447
Male secondary education (lagged)	0.001	-0.350	0.232*	-0.250	0.116	-0.128	-1.215	1.703
	(0.098)	(0.546)	(0.133)	(1.108)	(0.226)	(0.340)	(1.235)	(18.137
Price level of investment (lagged)	-0.000	0.000	0.000	-0.000	-0.000	-0.000	0.000	0.007
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.073)
Constant	0.025	0.808***	0.061		0.015	0.144		
	(0.061)	(0.238)	(0.070)		(0.136)	(0.329)		

Conclusions (so to speak…)

1. New evidence on the cross-country association between inequality and subsequent growth.
   - Greatly improved comparability in inequality measures.
   - More recent panel, with more diverse coverage
   - Incorporates recent innovations in GMM estimation, as well as concerns with its robustness

2. Overall, the evidence is (weakly) suggestive of a negative relationship between total inequality and growth, unlike previous panel results.
   - A positive relationship appears much less likely than in Forbes (2000).

3. But the evidence in support of our original hypothesis on the link between inequality of opportunity and growth is weak and unstable.
   - In contrast to within-country estimates, such as Marrero and Rodrigues (2013)
   - Measurement error, particular in the lower-bound estimate for I.Op. could be likely to blame.
   - But other explanations – such as no relationship – cannot be excluded.
   - Until much better measurement of circumstances is available over time and across countries, cross-country regressions – however, carefully executed – are unlikely to be able to teach us much.