

The IDLA Dataset: a Tool to Analyze Recent Changes in Income Inequality in L.A. *(mark 0)*

by

Bruno Martorano (bruno.martorano@unifi.it)
and Giovanni Andrea Cornia (giovanniandrea.cornia@unifi.it)
University of Florence

1. Overall purpose and types of analyses that can be carried out on IDLA

The dataset on Income Distribution in Latin America (IDLA) has been developed in the context of the UNU – WIDER’s Research Project on “The New Policy Model, Inequality and Poverty in Latin America: Evidence from the Last Decade and Prospects for the Future” coordinated by Giovanni Andrea Cornia. (http://www.wider.unu.edu/research/current-programme/en_GB/Impact-of-Economic-Crisis/). It is expected that the IDLA dataset will be updated late in 2011.

Such dataset is accessible in an open source mode to stimulate research by scholars with interests in the topic of income inequality in the region. Authors are kindly requested to refer to the database whenever they make use of it.

The IDLA database compiles published¹ statistical information useful for the analysis of income distribution in the Latin America region over the last two decades, a period marked by important changes in the field of income inequality, as well as in those of macroeconomic, taxation, labour market, and social policy in most countries of the region. In addition to this introduction (which lists the variables included in the database, their definitions and their sources), the IDLA dataset comprises an excel file divided into three sub-files focussing on economic, political and social variables. A second excel file provides a comparison between trends in the Gini coefficients of income inequality derived from different data sources.

The economic data are grouped according to the main issue to which they relate, i.e.: (i) economic structure and conditions; (ii) fiscal policies; (iii) monetary policies; (iv) external conditions. In turn, the political data are grouped according to the following categories: (i) institutional framework – i.e. electoral rules, forms of government, federalism – and (ii) political conditions – i.e. corruption, bureaucratic quality, and democracy. Lastly, the data on social conditions are clustered into the following groups: (i) income inequality; (ii) poverty; (iii) human capital; (iv) labour market conditions; (v) social fractionalization.

With respect to its possible use, the IDLA database is particularly suited for macro panel analyses of Latin America as a whole or of some of its sub-regions. Given the relatively short number of yearly observations (about 20), the IDLA database is less useful for single country time-series analyses.

As for the specific topics which could be analyzed, the information included in the IDLA dataset allows to analyze the inequality impact of several policies over the last two decades. For example,

¹ In very few cases the data were interpolated. In such a case the data are indicated with a brown-brick color.

Cornia (2010) and Cornia and Martorano (2010) analyzed the factors explaining the changes in income inequality between 1990 and 2007 in Latin America. Particular attention was paid to the impact of macroeconomic, educational and social policy changes introduced during this period and to the impact of democratization and shift of political regimes towards the centre-left. The data reported in the IDLA dataset allowed to test econometrically the importance of these and other factors on data for 18 countries and the period 1990- 2007.

The IDLA dataset allows also to investigate which factors explain the radical changes in taxation and fiscal policy which have occurred in Latin America during the last decade, as well as their impact on equity and growth. For example, Martorano (2011) tried to disentangle which historical, cultural and socio-economic factors explain the recent rise in tax/GDP ratios in Latin America. In addition, he investigated whether the recent increase in taxation contributed to macroeconomic stabilization and growth. Finally, using the same dataset Martorano (2011) analysed the impact of fiscal policy changes on income inequality.

Finally, the database contains social and political information which can be useful to investigate the factors behind the changes in the social and political landscape.

The IDLA has already been used for the preparation of the following research works:

- Cornia, G. A. (2010), “Income Distribution under Latin America's New Left Regimes”, *Journal of Human Development and Capabilities*, Taylor and Francis Journals, vol. 11(1), pages 85-114.

- Cornia, G. A. and B. Martorano (2010), “Policies for Reducing Income Inequality: Latin America during the Last Decade”, Social and Economic Policy Working Papers, UNICEF.

[http://www.unicef.org/socialpolicy/files/Postscript Formatted Policies for reducing income inequality.pdf](http://www.unicef.org/socialpolicy/files/Postscript_Formatted_Policies_for_reducing_income_inequality.pdf)

- Martorano, B. (2010), “Education Inequality and Wage Inequality in Latin America”, term paper written as part of the PhD program in Development Economics, University of Florence, mimeo, (available on request from the author).

- Cornia, G. A. and B. Martorano (2011), “External shocks, policy changes, and income distribution: Latin America during the last decade” in Edmund Valpy Fitzgerald, Judith Heyer and Rosemary Thorpe (eds.) “Overcoming the Persistence of Inequality and Poverty: Essays in honor of Frances Stewart”, MacMillan Palgrave, London, (March 2011), ISBN 978-0-230-24970-7.

- Martorano, B. (2011), “The Great Tax Transformation”, PhD Thesis in Development Economics, University of Florence, mimeo, (available on request from the author).

2. Time period and countries covered

The IDLA dataset includes information on 18 countries. Of these 10 are from South America (Argentina, Bolivia, Brazil, Chile, Colombia, Ecuador, Paraguay, Peru, Uruguay, Venezuela) and eight from Central America or the Caribbean (Costa Rica, Dominican Republic, El Salvador, Guatemala, Honduras, Mexico, Nicaragua, Panama). The time span covered is 1990-2008, with annual information with the exception of a few variables. For example, the data on urban population and illiteracy rate of the population are provided for five-years periods, i.e. for “1990”, “1995”, “2000”, “2005”. Also for ethnic, linguistic and religious fractionalization variables, only one

observation is available for the entire period. Finally, country and currency names are string variables.

3. Variables description and sources:

The dataset includes **137 variables** on economic (60), political - institutional (44) and social (33) factors which are generally used for the analysis of inequality changes.

The vast majority of the data included in the IDLA derive from existing published sources (e. g. international published sources, and national sources). As for other developing regions, it is not always easy to find reliable data. All together, the database includes 342 (18x19) cells for each variable, though missing data (about 8% of total sample) reduce the number of data usable data-strings (for more details see Appendix 1).

As noted, in the case of the Gini coefficient of income distribution (the most important variable for the analyses of distributive changes), the IDLA dataset provides a comparison of the trends of Gini data originating from different sources (see Appendix 2)². The main source of information for this variable is the SEDLAC database which computes the Gini coefficients following a standardized approach on the basis of micro data from household surveys. Although the frequency of household surveys increased since the early 1990s, in several years such precious source of information is not available. To deal with the ensuing problem of missing data, information coming from other data sources was also used, while some data (21%) were interpolated when there were one or two missing years in a clearly trended time series. This allowed to reduce the share of missing information to 6%. The interpolated data are clearly identified in the database by a different colour. This approach has the advantage of enlarging the size of the panel of observations (for more details see appendix B).

Often, the researcher has to face important problems in terms of differences across surveys in terms of the 'income concept' and geographical coverage used. Differences are most common between countries, but they tend to arise also from changes over time within countries (see Wiid2b documentation, <http://62.237.131.23/wiid/WIID2c.pdf>).

For this reason, IDLA uses only information related to inequality estimates based on income. A second question is related to what kind of income definition was used. In the last version of WIID, the Gini coefficients for the Latin American countries were calculated on both net income and gross income. In particular, in almost all LAC countries the surveys no questions are asked on taxes paid. For this reason, it is at times impossible to know whether a given Gini coefficient is computed on the basis of gross or net income data. It is likely that formal sector wage earners mostly report net income, and the few self-employed who pay direct taxes report the gross income.

The list of variables includes:

3.1 Economic Data

1) **year** = year

2) **country** = country name

² The same procedure is used to compute the informal sector size. The vast majority of the data is taken from international published sources and - in only few cases – the data originate from national sources. In very few cases (when 1 or 2 yearly data part of stable trend were missing) the data were interpolated, so as to increase the number of degrees of freedom in regression analysis.

3) **gdpc_ppp** = Per capita Gross Domestic Product, PPP (constant 2005 international dollar). PPP GDP is gross domestic product converted to international dollars using purchasing power parity rates. An international dollar has the same purchasing power over GDP as the U.S. dollar has in the United States. GDP at purchaser's prices is the sum of gross value added by all resident producers in the economy plus any product taxes and minus any subsidies not included in the value of the products. It is calculated without making deductions for depreciation of fabricated assets or for depletion and degradation of natural resources. Data are in constant 2005 international dollars.

Source: WDI

4) **gdpc_2005** = Per capita Gross Domestic Product at constant market prices (dollar of 2005).

Source: ERS International Macroeconomic Data Set

5) **gdpc_2000** = Per Capita Gross Domestic Product at constant market prices (dollars of 2000).

Source: CEPALSTAT

6) **agr** = Agriculture Value Added Industry (as a percentage of GDP).

Source: WDI

7) **ind** = Industry Value Added (as a percentage of GDP).

Source: WDI

8) **ma** = Manufacturing Value Added (as a percentage of GDP).

Source: WDI

9) **investment** = Gross capital formation (% of GDP). Gross capital formation (formerly gross domestic investment) consists of outlays on additions to the fixed assets of the economy plus net changes in the level of inventories. Fixed assets include land improvements (fences, ditches, drains, and so on); plant, machinery, and equipment purchases; and the construction of roads, railways, and the like, including schools, offices, hospitals, private residential dwellings, and commercial and industrial buildings. Inventories are stocks of goods held by firms to meet temporary or unexpected fluctuations in production or sales, and "work in progress." According to the 1993 SNA, net acquisitions of valuables are also considered capital formation.

Source: WDI

10) **consumption** = Household final consumption expenditure, etc. (% of GDP). Household final consumption expenditure (formerly private consumption) is the market value of all goods and services, including durable products (such as cars, washing machines, and home computers), purchased by households. It excludes purchases of dwellings but includes imputed rent for owner-occupied dwellings. It also includes payments and fees to governments to obtain permits and licenses. Here, household consumption expenditure includes the expenditures of nonprofit institutions serving households, even when reported separately by the country. This item also includes any statistical discrepancy in the use of resources relative to the supply of resources.

Source: WDI

11) **pub_exp** = General government final consumption expenditure (% of GDP). General government final consumption expenditure (formerly general government consumption) includes all government current expenditures for purchases of goods and services (including compensation of employees). It also includes most expenditures on national defense and security, but excludes government military expenditures that are part of government capital formation.

Source: WDI

12) **SOC_tot** = Social public expenditure as percentage of the Gross Domestic Product. It is the estimate of the allocation of public resources for expenditure in social sectors, as a percentage of total GDP.

Source: CEPALSTAT

13) **SOC_edu** = Education public expenditure as percentage of the Gross Domestic Product. It is the estimate of the allocation of public resources for expenditure in education, as a percentage of total GDP.

Source: CEPALSTAT

14) **SOC_health** = Health public expenditure as percentage of the Gross Domestic Product. It is the estimate of the allocation of public resources for expenditure in health, as a percentage of total GDP.

Source: CEPALSTAT

15) **SOC_ssec** = Social security public expenditure as percentage of the Gross Domestic Product. It is the estimate of the allocation of public resources for expenditure in social security, as a percentage of total GDP.

Source: CEPALSTAT

16) **SOC_house** = Housing public expenditure as percentage of the Gross Domestic Product. It is the estimate of the allocation of public resources for expenditure in housing, as a percentage of total GDP.

Source: CEPALSTAT

17) **t1** = Tax revenue as percentage of GDP. It is composed of compulsory transfers to the (central or general) government sector. Tax revenue is referred to central governments and thus it excludes taxes imposed by subnational governments. For Argentina, Bolivia, Brazil, Chile, Colombia and Costa Rica we have provided more comprehensive data for general rather than just central government.

Source: CEPALSTAT

18) **t2** = Direct Tax Revenue as percentage of GDP. It is referred to central governments and thus it excludes taxes imposed by subnational governments. For Argentina, Bolivia, Brazil, Chile, Colombia and Costa Rica we have provided more comprehensive data for general rather than just central government.

Source: CEPALSTAT

19) **t3** = Taxes on income, profits and capital gains (as percentage of GDP). ECLAC reports that taxes on income, profits, and capital gains generally are levied on (i) wages, salaries, tips, fees, commissions, fringe benefits, and other compensation for labor services; (ii) interest, dividends, rent, and royalty incomes; (iii) capital gains and losses, including capital gain distributions of investment funds; (iv) profits of corporations and trusts; (v) taxable portions of social security, pension, annuity, life insurance, and other retirement account distributions; and (vi) miscellaneous other income items. It is referred to central governments and thus it excludes taxes imposed by subnational governments. For Argentina, Bolivia, Brazil, Chile, Colombia and Costa Rica we have provided more comprehensive data for general rather than just central government.

Source: CEPALSTAT

20) **t4** = Taxes on individuals as percentage of GDP. It is referred to central governments and thus it excludes taxes imposed by subnational governments. For Argentina, Bolivia, Brazil, Chile, Colombia and Costa Rica we have provided more comprehensive data for general rather than just central government.

Source: CEPALSTAT

21) **t5** = Taxes on corporations and enterprises as percentage of GDP. It is referred to central governments and thus it excludes taxes imposed by subnational governments. For Argentina, Bolivia, Brazil, Chile, Colombia and Costa Rica we have provided more comprehensive data for general rather than just central government.

Source: CEPALSTAT

22) **t6** = Unallocable taxes as percentage of GDP. ECLAC notes that taxes are treated as unallocable when the information needed to determine whether taxes should be attributed to either of these categories is not available. It is referred to central governments and thus it excludes taxes imposed by subnational governments. For Argentina, Bolivia, Brazil, Chile, Colombia and Costa Rica we have provided more comprehensive data for general rather than just central government.

Source: CEPALSTAT

23) **t7** = Taxes on property as percentage of GDP. Taxes on property are levied on the use, ownership, or transfer of wealth and are divided into six categories: recurrent taxes on immovable property; recurrent taxes on net wealth; estate, inheritance, and gift taxes; taxes on financial and capital transactions; other nonrecurrent taxes on property; and other recurrent taxes on property. It is referred to central governments and thus it excludes taxes imposed by subnational governments. For Argentina, Bolivia, Brazil, Chile, Colombia and Costa Rica we have provided more comprehensive data for general rather than just central government.

Source: CEPALSTAT

24) **t8** = Other direct taxes as percentage of GDP. It is referred to central governments and thus it excludes taxes imposed by subnational governments. For Argentina, Bolivia, Brazil, Chile, Colombia and Costa Rica we have provided more comprehensive data for general rather than just central government.

Source: CEPALSTAT

25) **t9** = Indirect Tax Revenue as percentage of GDP. It is referred to central governments and thus it excludes taxes imposed by subnational governments. For Argentina, Bolivia, Brazil, Chile, Colombia and Costa Rica we have provided more comprehensive data for general rather than just central government.

Source: CEPALSTAT

26) **t10** = General taxes on goods and services as percentage of GDP. Taxes on goods and services include value-added taxes; general sales taxes; single-stage and cumulative multistage taxes, where “stage” refers to stage of production or distribution; excises; taxes levied on the use of motor vehicles or other goods; taxes levied on permission to use goods or perform certain activities; taxes on the extraction, processing, or production of minerals and other products. It is referred to central governments and thus it excludes taxes imposed by subnational governments. For Argentina, Bolivia, Brazil, Chile, Colombia and Costa Rica we have provided more comprehensive data for general rather than just central government.

Source: CEPALSTAT

27) **t11** = Taxes on specific goods and services as percentage of GDP. It is referred to central governments and thus it excludes taxes imposed by subnational governments. For Argentina, Bolivia, Brazil, Chile, Colombia and Costa Rica we have provided more comprehensive data for general rather than just central government.

Source: CEPALSTAT

28) **t12** = Taxes on international trade and transactions as percentage of GDP. It includes customs and other import duties and taxes on exports. It is referred to central governments and thus it excludes taxes imposed by subnational governments. For Argentina, Bolivia, Brazil, Chile, Colombia and Costa Rica we have provided more comprehensive data for general rather than just central government.

Source: CEPALSTAT

29) **t13** = Other indirect taxes as percentage of GDP. It is referred to central governments and thus it excludes taxes imposed by subnational governments. For Argentina, Bolivia, Brazil, Chile, Colombia and Costa Rica we have provided more comprehensive data for general rather than just central government.

Source: CEPALSTAT

30) **t14** = Other taxes as percentage of GDP. ECLAC reports that this item covers revenue from taxes levied predominantly on a base or bases other than those described under the preceding tax headings. Also the item includes taxes on persons that are not based on income or presumptive income (poll taxes, capitation taxes or head taxes), stamp taxes that do not fall exclusively or predominantly on a single class of transactions or activities covered by other taxes. It is referred to central governments and thus it excludes taxes imposed by subnational governments. For Argentina, Bolivia, Brazil, Chile, Colombia and Costa Rica we have provided more comprehensive data for general rather than just central government.

Source: CEPALSTAT

31) **t15** = Social contributions as percentage of GDP. They are actual or imputed receipts from either employers on behalf of their employees or from employees, self-employed, or nonemployed persons on their

own behalf that secure entitlement to social benefits for their contributors, their dependents, or their survivors. The contributions may be compulsory or voluntary. It is referred to central governments and thus it excludes taxes imposed by subnational governments. For Argentina, Bolivia, Brazil, Chile, Colombia and Costa Rica we have provided more comprehensive data for general rather than just central government.

Source: CEPALSTAT

32) **tax** = Tax revenue (including social contributions) as percentage of GDP. It is referred to central governments and thus it excludes taxes imposed by subnational governments. For Argentina, Bolivia, Brazil, Chile, Colombia and Costa Rica we have provided more comprehensive data for general rather than just central government.

Source: CEPALSTAT

33) **n_tax** = Nontax revenue as percentage of GDP.

Source: CEPALSTAT

34) **balance** = Overall fiscal balance as percentage of GDP.

Source: CEPALSTAT

35) **import** = Imports of goods and services (% of GDP). Imports of goods and services represent the value of all goods and other market services received from the rest of the world. They include the value of merchandise, freight, insurance, transport, travel, royalties, license fees, and other services, such as communication, construction, financial, information, business, personal, and government services. They exclude compensation of employees and investment income (formerly called factor services) and transfer payments.

Source: WDI

36) **export** = Exports of goods and services (% of GDP). Exports of goods and services represent the value of all goods and other market services provided to the rest of the world. They include the value of merchandise, freight, insurance, transport, travel, royalties, license fees, and other services, such as communication, construction, financial, information, business, personal, and government services. They exclude compensation of employees and investment income (formerly called factor services) and transfer payments.

Source: WDI

37) **rer** = Real Exchange Rate (US Ag. Trade Weighted Exchange Rate, 2005=100).

Source: ERS

38) **reer** = Index of Real Effective Exchange Rate (2000 = 100). Real effective exchange rate is the nominal effective exchange rate (a measure of the value of a currency against a weighted average of several foreign currencies) divided by a price deflator or index of costs.

Source: Economic Survey of Latin America and the Caribbean (several issues)

39) **err** = exchange rate regime:

- (1) dollarization
- (2) currency board
- (3) other fixed parity regimes
- (4) crawling pegs and moving bands regimes
- (5) flexible regime

Source: Reinhart and Rogoff (2002); Economic Survey of Latin America and the Caribbean

40) **unna_cu** = Currency. Name of the currency used in the country.

Source: United Nations Statistics Divisions 2009 reported in Teorell et al (2010)

41) **d1** = Average applied import tariff rates on manufactured goods, ores and metals (Weighted average).

Source: UNCTAD

- 42) **dt2** = Average applied import tariff rates on ores and metals (Weighted average).
Source: UNCTAD
- 43) **dt3** = Average applied import tariff rates on Manufactured goods (Weighted average).
Source: UNCTAD
- 44) **dt4** = Average applied import tariff rates on Chemical products (Weighted average).
Source: UNCTAD
- 45) **dt5** = Average applied import tariff rates on machinery and transport equipment (Weighted average).
Source: UNCTAD
- 46) **dt6** = Average applied import tariff rates on other manufactured goods (Weighted average).
Source: UNCTAD
- 47) **tot1** = International terms of trade, fob (2000 = 100).
Source: CEPALSTAT
- 48) **tot2** = International terms of trade, services (2000 = 100).
Source: CEPALSTAT
- 49) **tot3** = International terms of trade, goods and services (2000 = 100).
Source: CEPALSTAT
- 50) **fdi_1** = Foreign direct investment: inward flows measured as percentage of GDP. UNCTAD reports that a direct investment enterprise is defined as an incorporated or unincorporated enterprise in which the direct investor, resident in another economy, owns 10 percent or more of the ordinary shares of voting power (or the equivalent). However, this criterion is not strictly observed by all countries reporting. For more detailed information on concepts presented in this table, please refer to the IMF Balance of Payments Manual (BPM5, 1993), UNCTAD World Investment Report 2007, and OECD Detailed Benchmark Definition of FDI.
Source: UNCTAD
- 51) **fdi_2** = Foreign direct investment: outward flows measured as percentage of GDP.
Source: UNCTAD
- 52) **fdi_3** = Foreign direct investment: inward stocks measured as percentage of GDP
Source: UNCTAD
- 53) **fdi_4** = Foreign direct investment: outward stocks measured as percentage of GDP.
Source: UNCTAD
- 54) **fdi** = Net Foreign direct investment stocks measured as percentage of GDP
Source: UNCTAD
- 55) **remittance** = Workers' remittances Receipts as percentage of GDP.
Source: USAID, UNCTAD and WDI
- 56) **aid** = Net ODA received (% of GNI). Net official development assistance is disbursement flows (net of repayment of principal) that meet the DAC definition of ODA and are made to countries and territories on the DAC list of aid recipients.
Source: WDI
- 57) **debt** = External debt as percentage of GDP.
Source: CEPALSTAT

58) **m2** = Money and quasi money (M2) as % of GDP. Money and quasi money comprise the sum of currency outside banks, demand deposits other than those of the central government, and the time, savings, and foreign currency deposits of resident sectors other than the central government. This definition of money supply is frequently called M2; it corresponds to lines 34 and 35 in the International Monetary Fund's (IMF) International Financial Statistics (IFS).

Source: WDI

59) **cpi** = Inflation measured by the average consumer price. Data for inflation are averages for the year, not end-of-period data. The index is based on annual percent change

Source: WEO

60) **rir** = Real interest rate (%). Real interest rate is the lending interest rate adjusted for inflation as measured by the GDP deflator.

Source: WDI

3.2 Political Data

1) **year** = year

2) **country** = country name

3) **polity2** = The polity score ranges from +10 (strongly democratic) to -10 (strongly autocratic). Democracy is conceived as three essential, interdependent elements: i) the presence of institutions and procedures through which citizens can express effective preferences about alternative policies and leaders; ii) the existence of institutionalized constraints on the exercise of power by the executive; iii) the guarantee of civil liberties to all citizens in their daily lives and in acts of political participation. Autocracy is defined in terms of the presence of a distinctive set of political characteristics as: restriction or suppression of competitive political participation; chief executives chosen in a regularized process of selection within the political elite, and once in office they exercise power with few institutional constraints, etc.

Source: Polity IV Project

4) **durable** = Regime Durability (years). The number of years since the most recent regime change or the end of transition period defined by the lack of stable political institutions.

Source: Polity IV Project

5) **chga_demo** = Democracy dummy (Coded 1 if democracy, 0 otherwise). A regime is considered a democracy if the executive and the legislature is directly or indirectly elected by popular vote, multiple parties are allowed, there is de facto existence of multiple parties outside of regime front, there are multiple parties within the legislature, and there has been no consolidation of incumbent advantage (e.g. unconstitutional closing of the lower house or extension of incumbent's term by postponing of subsequent elections). Transition years are coded as the regime that emerges in that year.

Source: Cheibub, et al (2009) reported in Teorell et al (2010)

6) **chga_hinst** = Regime Institutions. Six-fold classification of political regimes, coded:

(0) Parliamentary democracy

(1) Mixed (semi-presidential) democracy

(2) Presidential democracy

(3) Civilian dictatorship

(4) Military dictatorship

(5) Royal dictatorship

Source: Cheibub, et al (2009) reported in Teorell et al (2010)

7) **b_system** = Regime Type defined as: (0) Presidential; (1) Assembly - elected President and (2) Parliamentary.

Source: Database of Political Institutions 2010 (see Keefer 2010)

8) **b_erlc** = Party of Chief Executive. This variable is defined considering party orientation with respect to economic policy: (1) Right, for parties that are defined as conservative, Christian democratic, or rightwing; (2) Left, for parties that are defined as communist, socialist, social democratic, or left-wing; and (3) Center, for parties that are defined as centrist or when party position can best be described as centrist (e.g. the party advocates strengthening private enterprise in a social-liberal context).

Source: Database of Political Institutions 2010 (see Keefer 2010)

9) **b_gov** = Largest Government Party orientation: Right, Left or Center (see variable **b_erlc** for more information).

Source: Database of Political Institutions 2010 (see Keefer 2010)

10) **social_dem** = Dummy denoting a country with a social-democratic government.

Source: authors' elaboration

11) **populista** = Dummy denoting a country with a populista government.

Source: authors' elaboration

12) **b_yrsoffic** = Year in office of the chief executive.

Source: Database of Political Institutions 2010 (see Keefer 2010)

13) **b_pr** = Proportional Representation. Dummy variable, "1" if candidates are elected based on the percent of votes received by their party and/or if our sources specifically call the system "proportional representation"; "0" otherwise

Source: Database of Political Institutions 2010 (see Keefer 2010)

14) **b_cl** = Closed Lists. Dummy variable, 1 when proportional representation is used (**b_pr**) and voters cannot express preferences for candidates within a party list.

Source: Database of Political Institutions 2010 (see Keefer 2010)

15) **van_comp** = Vanhanen – Index of Competition which ranges from 0 (only one party received 100 % of votes) to 100 (each voter cast a vote for a distinct party). The competition variable portrays the electoral success of smaller parties, that is, the percentage of votes gained by the smaller parties in parliamentary and/or presidential elections. The variable is calculated by subtracting from 100 the percentage of votes won by the largest party (the party which wins most votes) in parliamentary elections or by the party of the successful candidate in presidential elections.

Source: Vanhanen. FSD1289, version 4.0 (2009-12-14).

16) **van_part** = Vanhanen – Index of Participation calculated as the percentage of the total population who actually voted in the election. In the case of indirect elections, only votes cast in the final election are taken into account. If electors have not been elected by citizens, only the number of actual electors is taken into account, which means that the degree of participation drops to the value 0. If an election to choose electors has been held, the participation variable is calculated from the number and distribution of votes in that election. National referendums raise the variable value by five percent and state (regional) referendums by one percent for the year they are held. Referendums can add the degree of participation at maximum by 30 percent a year. The value of the combined degree of participation cannot be higher than 70 percent, even in cases where the sum of participation and referendums would be higher than 70.

Source: Vanhanen. FSD1289, version 4.0 (2009-12-14).

17) **van_dem** = Vanhanen – Index of Democracy. The index of democratization is formed by multiplying the competition and the participation variables and then dividing the outcome by 100.

Source: Vanhanen. FSD1289, version 4.0 (2009-12-14).

18) **b_frac** = Total Fractionalization which measures the probability that two randomly chosen deputies in the legislature belong to different parties.

Source: Database of Political Institutions 2010 (see Keefer 2010)

19) **b_maj** = Majority Seats measured as the number of government seats divided by total seats in the legislature.

Source: Database of Political Institutions 2010 (see Keefer 2010)

20) **b_legelec** = Legislative Election index.

No legislature: 1

Unelected legislature: 2

Elected, 1 candidate: 3

1 party, multiple candidates: 4

multiple parties are legal but only one party won seats: 5

multiple parties DID win seats but the largest party

received more than 75% of the seats: 6

largest party got less than 75%: 7

Source: Database of Political Institutions 2010 (see Keefer 2010)

21) **b_exelec** = Executive Election index. Uses same scale as Legislative IEC

Source: Database of Political Institutions 2010 (see Keefer 2010)

22) **b_fraud** = Fraud or Candidate Intimidation Affection. Dummy variable, 1 when opposition is officially legal but reported vote fraud or candidate intimidation were serious enough to affect the outcome of elections. If not an election year, or if elected government has been deposed, records to the most recent election.

Source: Database of Political Institutions 2010 (see Keefer 2010)

23) **ciri_assn** = Freedom of Assembly and Association. Citizens' rights to freedom of assembly and association are:

(0) Severely restricted or denied completely to all citizens

(1) Limited for all citizens or severely restricted or denied for selected groups

(2) Virtually unrestricted and freely enjoyed by practically all citizens

Source: Cingranelli and Richards (2010) reported in Teorell et al (2010)

24) **ciri_polpris** = Political Imprisonment. Are there any people imprisoned because of their political, religious, or other beliefs?

(0) Yes, many

(1) Yes, but few

(2) None

Source: Cingranelli and Richards (2010) reported in Teorell et al (2010)

25) **iaep_bp** = Banned Parties. Equals 1 if there are banned parties, and 0 otherwise.

Source: Regan and Clark (2010) reported in Teorell et al (2010)

26) **iaep_cc** = Constitutional Court. Equals 1 if the country according to the constitution has a national constitutional court, and 0 otherwise. In some cases, a council with powers of a constitutional court may exist, though it may not be part of the formal judiciary. In such cases, this nonjudicial council with powers of a constitutional court is coded as the constitutional court.

Source: Regan and Clark (2010) reported in Teorell et al (2010)

27) **iaep_ufs** = Unitary or Federal State

(1) Unitary system

(2) Confederation

(3) Federal system

Source: Regan and Clark (2010) reported in Teorell et al (2010)

28) **f_overall** = Overall Freedom Index (1 is free, 2 is partially free, and 3 represents not free). As defined by Freedom House, freedom encompasses two sets of characteristics grouped under political rights and civil liberties.

Source: USAID

29) **f_civil** = The Civil Liberties Index is a scaled index of the Civil Liberties Aggregate Score so that 1 is the highest, and 7 is the lowest measure of freedom. Civil liberties are the freedoms to develop views, institutions, and personal autonomy apart from the state.

Source: USAID

30) **f_politic** = The Political Rights Index is a scaled index of the Political Rights Aggregate Score so that 1 is the highest, and 7 is the lowest measure of freedom. Political rights enable people to participate freely in the political process. In this case, political process refers to the system by which the polity chooses the authoritative policy makers and attempts to make binding decisions affecting the national, regional, or local community. A system is genuinely free or democratic to the extent that the people have a choice in determining the nature of the system and its leaders.

Source: USAID

31) **dr_eg** = Economic Globalization Index (between 0 and 100, where higher values indicate a higher degree of globalization). Economic globalization is here defined as the long distance flows of goods, capital and services as well as information and perceptions that accompany market exchanges. It is measured by actual flows of trade and investments, and by restrictions on trade and capital such as tariff rates.

Source: Dreher (2006); Dreher et al (2008) reported in Teorell et al (2010)

32) **dr_ig** = Index of Globalization (between 0 and 100, where higher values indicate a higher degree of globalization). The overall index of globalization is the weighted average of the following variables: economic globalization, social globalization and political globalization (**dr_eg**, **dr_sg** and **dr_pg**). Most weight has been given to economic followed by social globalization.

Source: Dreher (2006); Dreher et al (2008) reported in Teorell et al (2010)

33) **dr_pg** = Political Globalization (between 0 and 100, where higher values indicate a higher degree of globalization). Political globalization is measured by the number of embassies and high commissions in a country, the number of international organizations of which the country is a member, the number of UN peace missions the country has participated in, and the number of international treaties that the country has signed since 1945.

Source: Dreher (2006); Dreher et al (2008) reported in Teorell et al (2010)

34) **dr_sg** = Social Globalization (between 0 and 100, where higher values indicate a higher degree of globalization). Social globalization is measured by three categories of indicators. The first is personal contacts, such as telephone traffic and tourism. The second is information flows, e.g. number of Internet users. The third is cultural proximity, e.g. trade in books and number of Ikea warehouses per capita.

Source: Dreher (2006); Dreher et al (2008) reported in Teorell et al (2010)

35) **icrg_qog** = ICRG indicator of Quality of Government scaled 0-1. Higher values indicate higher quality of government.

Source: International Country Risk Guide – The PRS Group, reported in Teorell et al (2010)

36) **wgi_1** = “Voice and Accountability” lies between –2.5 and 2.5 with higher scores corresponding to better outcomes. It measures the extent to which a country’s citizens are able to participate in selecting their government, as well as freedom of expression, freedom of association, and a free media.

Source: Worldwide Governance Indicators

37) **wgi_2** = “Political Stability” lies between –2.5 and 2.5 with higher scores corresponding to better outcomes. It combines several indicators which measure perceptions of the likelihood that the government in

power will be destabilized or overthrown by possibly unconstitutional and/or violent means, including domestic violence and terrorism.

Source: Worldwide Governance Indicators

38) **wgi_3** = Government Effectiveness lies between –2.5 and 2.5 with higher scores corresponding to better outcomes. It measures the quality of public services, the capacity of the civil service and its independence from political pressures; and the quality of policy formulation.

Source: Worldwide Governance Indicators

39) **wgi_4** = Regulatory Quality lies between –2.5 and 2.5 with higher scores corresponding to better outcomes. It measures the ability of the government to provide sound policies and regulations that enable and promote private sector development.

Source: Worldwide Governance Indicators

40) **wgi_5** = Rule of Law lies between –2.5 and 2.5 with higher scores corresponding to better outcomes. It measures the extent to which agents have confidence in and abide by the rules of society. These include perceptions of the incidence of crime, the effectiveness and predictability of the judiciary, and the enforceability of contracts.

Source: Worldwide Governance Indicators

41) **wgi_6** = Control of Corruption lies between –2.5 and 2.5 with higher scores corresponding to better outcomes. It measures the extent to which public power is exercised for private gain, including both petty and grand forms of corruption, as well as “capture” of the state by elites and private interests.

Source: Worldwide Governance Indicators

42) **trust_1** = trustness in congress (%). Percentage of people which trust in congress.

Source: CEPALSTAT

43) **trust_2** = trustness in law (%). Percentage of people which trust in congress.

Source: CEPALSTAT

44) **trust_3** = trustness in polity (%). Percentage of people which trust in congress.

Source: CEPALSTAT

3.3 Social Data

1) **year** = year

2) **country** = country name

3) **gini** = Gini index on income. The Gini coefficients for Latin American countries are calculated on a mixture of net income and gross income concept. As reported above, the surveys in almost all LAC countries there are no questions on taxes, so as suggested by Gasparini, we really don't know whether it's gross or net income. We believe that formal wage earners mostly report the net income, and the few self-employed who pay direct taxes report the gross income.

Source: SEDLAC, WIID, CEPALSTAT, WDI, SWIID Version 3.0 and national source. Some missing data were obtained by interpolation (see Appendix 2).

4) **pov1** = Poverty gap at \$1.25 a day (PPP) (%). Poverty gap is the mean shortfall from the poverty line (counting the nonpoor as having zero shortfall), expressed as a percentage of the poverty line. This measure reflects the depth of poverty as well as its incidence.

Source: WDI

5) **pov2** = Poverty gap at \$2 a day (PPP) (%).

Source: WDI

6) **pov3** = Poverty headcount ratio at \$1.25 a day (PPP) (% of population). Population below \$1.25 a day is the percentage of the population living on less than \$1.25 a day at 2005 international prices. As a result of revisions in PPP exchange rates, poverty rates for individual countries cannot be compared with poverty rates reported in earlier editions.

Source: WDI

7) **pov4** = Poverty headcount ratio at \$2 a day (PPP) (% of population)

Source: WDI

8) **pop** = Population, total. Total population is based on the de facto definition of population, which counts all residents regardless of legal status or citizenship--except for refugees not permanently settled in the country of asylum, who are generally considered part of the population of their country of origin. The values shown are midyear estimates.

Source: WDI

9) **d_pop** = Population density (people per sq. km of land area). Population density is midyear population divided by land area in square kilometers. Population is based on the de facto definition of population, which counts all residents regardless of legal status or citizenship--except for refugees not permanently settled in the country of asylum, who are generally considered part of the population of their country of origin. Land area is a country's total area, excluding area under inland water bodies, national claims to continental shelf, and exclusive economic zones. In most cases the definition of inland water bodies includes major rivers and lakes.

Source: WDI

10) **urban** = Urban population (% of total). Urban population refers to people living in urban areas as defined by national statistical offices. It is calculated using World Bank population estimates and urban ratios from the United Nations World Urbanization Prospects.

Source: WDI

11) **life** = Life expectancy at birth, total (years). Life expectancy at birth indicates the number of years a newborn infant would live if prevailing patterns of mortality at the time of its birth were to stay the same throughout its life.

Source: WDI

12) **umr** = Mortality rate, under-5 (per 1,000). Under-five mortality rate is the probability per 1,000 that a newborn baby will die before reaching age five, if subject to current age-specific mortality rates.

Source: WDI

13) **mmr** = Maternal mortality ratio (modeled estimate, per 100,000 live births). Maternal mortality ratio is the number of women who die during pregnancy and childbirth, per 100,000 live births. The data are estimated with a regression model using information on fertility, birth attendants, and HIV prevalence.

Source: WDI

14) **fertility** = Fertility rate, total (births per woman). Total fertility rate represents the number of children that would be born to a woman if she were to live to the end of her childbearing years and bear children in accordance with current age-specific fertility rates.

Source: WDI

15) **ily** = Illiteracy rate of population aged 15 years and over (%).

Source: CEPALSTAT

16) **hc_low** = share of adults aged 25-65 with 0 to 8 years of formal education. Given that the educational systems differ across countries and over time, SEDLAC includes people with different educational levels

(e.g. primary incomplete and complete) into the same category. SEDLAC considers that an acceptable cost to be paid in order to present simple statistics of educational levels in a similar format for all countries/years.
Source: SEDLAC

17) **hc_medium** = share of adults aged 25-65 with 9 to 13 years of formal education.
Source: SEDLAC

18) **hc_high** = share of adults aged 25-65 with more than 13 years of formal education.
Source: SEDLAC

19) **edu_years** = Average years of formal education in the adult population (aged 25 to 65). Most surveys record years of formal education. In those cases where this variable does not exist it is possible to approximate it from the maximum educational level attained by the person and her age. For some countries the distribution is truncated in the upper tail, since surveys do not ask on years of graduate education (e.g. Argentina).
Source: SEDLAC

20) **edu_gini** = Education Gini Index. It is computed as one measure of inequality in education. The use of educational-Ginis has been increasing in the last years.
Source: SEDLAC

21) **enr_i** = Primary school, net enrollment rates (%). Share of children in primary school age attending primary school.
Source: SEDLAC

22) **enr_ii** = Secondary school, net enrollment rates (%). Share of children in secondary school age attending secondary school.
Source: SEDLAC

23) **enr_iii** = Tertiary school net, enrollment rates (%). Share of children in tertiary school age attending tertiary school.
Source: SEDLAC

24) **mobile** = mobile phone users (per 100 people).
Source: WDI

25) **internet** = internet users (per 100 people).
Source: WDI

26) **informal** = Informal Sector (%). Urban population employed in jobs characterized by insecurity with respect to salaries, duration, social security, and so on. A person is defined as employed in the low productivity sector (informal) if she is an employer or salaried employee (professional, technical or neither) who works for an establishment employing up to five persons (micro enterprise), who works in a domestic employment, or who is an independent non-qualified worker (own-account and unpaid family workers with neither professional nor technical qualification). The index is measured as Percentage of total urban employed population. Some missing data are obtained by interpolation.
Source: CEPALSTAT, Labour Overview (ILO), SEDLAC and National Data

27) **mw** = Index of nominal minimum wages deflated by countries' CPI (2000 = 100). The indicator corresponds to the minimum wages for the formal sector.
Source: CEPALSTAT

28) **un** = Unemployment, total (% of total labor force). Unemployment refers to the share of the labor force that is without work but available for and seeking employment. Definitions of labor force and unemployment differ by country.
Source: WDI

29) **lab** = labour force participation rate. The index represents the percentage of the total population aged 15 years and over.

Source: CEPALSTAT

30) **wlab** = Labor force, female (% of total labor force). Female labor force as a percentage of the total show the extent to which women are active in the labor force. Labor force comprises people ages 15 and older who meet the International Labour Organization's definition of the economically active population.

Source: WDI

31) **al_ethnic** = Ethnic fractionalization. Reflects probability that two randomly selected people from a given country will not belong to the same ethnolinguistic group. The higher the number, the more fractionalized society. The definition of ethnicity involves a combination of racial and linguistic characteristics.

Source: Alesina et al (2003) reported in Teorell et al (2010)

32) **al_language** = Linguistic fractionalization. Reflects probability that two randomly selected people from a given country will not belong to the same linguistic group. The higher the number, the more fractionalized society.

Source: Alesina et al (2003) reported in Teorell et al (2010)

33) **al_religion** = Religious fractionalization. Reflects probability that two randomly selected people from a given country will not belong to the same religious group. The higher the number, the more fractionalized society.

Source: Alesina et al (2003) reported in Teorell et al (2010)

APPENDIX 1

Table 1a. Synthetic Table of Economic variables

Variable	Obs	Mean	Std. Dev.	Min	Max
<i>gdpc_ppp</i>	342	6663.65	2917.33	1742.94	13433.61
<i>gdpc_2005</i>	342	3334.17	1845.22	706.56	8378.35
<i>gdpc_2000</i>	342	3212.50	2039.93	632.70	9884.88
<i>agr</i>	318	11.36	5.83	3.68	27.78
<i>ind</i>	338	30.53	7.48	13.72	60.56
<i>ma</i>	342	18.43	4.45	2.20	28.34
<i>investment</i>	342	20.85	4.94	10.22	38.35
<i>consumption</i>	342	71.30	9.90	46.75	98.57
<i>pub_exp</i>	342	11.83	3.95	2.98	43.48
<i>soc_tot</i>	322	11.40	5.39	2.86	26.05
<i>soc_edu</i>	322	3.63	1.25	0.83	7.78
<i>soc_health</i>	322	2.52	1.27	0.22	5.82
<i>soc_ssec</i>	303	4.40	3.87	0.03	15.38
<i>soc_house</i>	322	1.10	0.77	0.00	3.11
<i>t1</i>	342	13.38	3.86	6.55	25.93
<i>t2</i>	342	4.16	2.17	0.76	15.67
<i>t3</i>	342	3.52	1.84	0.64	15.61
<i>t4</i>	196	0.65	0.60	0.00	2.18
<i>t5</i>	216	2.40	1.76	0.36	14.27
<i>t6</i>	335	1.55	1.49	-0.60	5.79
<i>t7</i>	342	0.60	0.73	-0.03	3.23
<i>t8</i>	342	0.04	0.12	0.00	0.62
<i>t9</i>	342	8.88	2.80	2.07	16.75
<i>t10</i>	342	5.25	2.63	0.00	12.97
<i>t11</i>	342	2.02	1.17	-1.35	6.91
<i>t12</i>	342	1.58	0.88	0.29	4.78
<i>t13</i>	342	0.03	0.08	-0.03	0.43
<i>t14</i>	342	0.34	0.47	-0.21	4.55
<i>t15</i>	337	2.79	2.34	0.00	9.02
<i>tax</i>	342	16.13	5.20	7.12	34.42
<i>tot1</i>	342	99.70	20.63	46.99	249.47
<i>tot2</i>	342	102.45	24.28	55.55	264.73
<i>tot3</i>	342	100.00	19.41	52.89	242.32
<i>tax</i>	342	16.13	5.20	7.12	34.42
<i>n_tax</i>	323	3.66	3.03	0.07	14.00
<i>balance</i>	334	-1.59	2.35	-11.54	8.82

Table 1b. Synthetic Table of Economic variables

Variable	Obs	Mean	Std. Dev.	Min	Max
<i>import</i>	342	34.46	18.50	4.63	98.07
<i>export</i>	342	30.47	17.16	6.57	100.70
<i>rer</i>	342	454.62	1188.30	0.81	6743.91
<i>reer</i>	323	106.17	24.78	60.30	237.10
<i>err</i>	306	3.86	1.17	1.00	5.00
<i>unna_cu</i>	0				
<i>d1</i>	189	9.38	4.01	2.16	26.95
<i>d2</i>	189	5.56	3.44	0.13	15.02
<i>d3</i>	189	9.49	4.14	2.25	28.70
<i>d4</i>	189	7.00	3.42	0.76	21.45
<i>d5</i>	189	9.62	4.89	1.57	32.88
<i>d6</i>	189	10.86	4.18	2.33	28.01
<i>tot1</i>	342	99.70	20.63	46.99	249.47
<i>tot2</i>	342	102.45	24.28	55.55	264.73
<i>tot3</i>	342	100.00	19.41	52.89	242.32
<i>fdi_1</i>	342	3.14	2.54	-3.01	14.58
<i>fdi_2</i>	335	0.86	2.99	-8.64	25.08
<i>fdi_3</i>	334	23.78	16.26	3.08	83.11
<i>fdi_4</i>	334	8.08	22.57	0.00	130.24
<i>fdi</i>	342	15.33	19.24	-65.49	82.67
<i>remittance</i>	333	3.24	4.38	0.00	21.90
<i>aid</i>	342	2.88	6.89	-0.73	72.06
<i>debt</i>	330	61.45	89.95	12.30	1064.40
<i>m2</i>	341	32.36	13.92	6.21	82.44
<i>cpi</i>	342	78.12	504.01	-1.17	7481.69
<i>rir</i>	318	13.32	18.34	-97.62	86.98

Table 2. Synthetic Table of Political variables

Variable	Obs	Mean	Std. Dev.	Min	Max
<i>polity2</i>	342	7.53	2.03	-3.00	10.00
<i>durable</i>	342	17.47	18.46	0.00	89.00
<i>chga_demo</i>	342	0.93	0.25	0.00	1.00
<i>chga_hinst</i>	342	2.07	0.25	2.00	3.00
<i>b_system</i>	342	0.01	0.08	0.00	1.00
<i>b_erc</i>	292	1.61	0.86	1.00	3.00
<i>b_gov</i>	307	1.61	0.82	1.00	3.00
<i>social_dem</i>	342	0.18	0.39	0.00	1.00
<i>populista</i>	342	0.07	0.26	0.00	1.00
<i>b_yrsoffic</i>	342	3.23	2.17	1.00	12.00
<i>b_pr</i>	342	0.94	0.23	0.00	1.00
<i>b_cl</i>	337	0.89	0.32	0.00	1.00
<i>van_comp</i>	342	50.43	10.59	0.00	70.00
<i>van_part</i>	342	36.65	12.54	0.00	69.00
<i>van_index</i>	342	18.77	7.67	0.00	45.90
<i>b_frac</i>	336	0.68	0.12	0.35	0.90
<i>b_maj</i>	339	0.52	0.16	0.09	1.00
<i>b_legelec</i>	342	0.26	0.44	0.00	1.00
<i>b_exelec</i>	342	0.22	0.41	0.00	1.00
<i>b_fraud</i>	342	0.11	0.31	0.00	1.00
<i>ciri_assn</i>	342	1.73	0.49	0.00	2.00
<i>ciri_polpris</i>	341	1.44	0.72	0.00	2.00
<i>iaep_bp</i>	288	0.33	0.47	0.00	1.00
<i>iaep_cc</i>	288	0.97	0.17	0.00	1.00
<i>iaep_ufs</i>	288	2.49	0.88	1.00	3.00
<i>f_overall</i>	342	1.50	0.50	1.00	2.00
<i>f_civil</i>	342	2.91	0.90	1.00	5.00
<i>f_politic</i>	342	2.49	1.02	1.00	6.00
<i>dr_eg</i>	306	57.08	10.30	29.73	85.48
<i>dr_ig</i>	306	57.13	8.61	36.85	75.11
<i>dr_pg</i>	306	68.51	13.81	39.87	93.12
<i>dr_sg</i>	306	50.41	10.48	23.79	70.95
<i>icrg_qog</i>	342	0.47	0.12	0.11	0.78
<i>wgi_1</i>	180	0.13	0.47	-0.62	1.21
<i>wgi_2</i>	180	-0.32	0.67	-2.14	0.97
<i>wgi_3</i>	180	-0.13	0.50	-1.01	1.39
<i>wgi_4</i>	180	0.02	0.62	-1.56	1.58
<i>wgi_5</i>	180	-0.43	0.62	-1.59	1.28
<i>wgi_6</i>	180	-0.30	0.65	-1.34	1.51
<i>trust_1</i>	234	27.58	10.86	5.90	58.50
<i>trust_2</i>	234	31.10	11.18	7.50	58.80
<i>trust_3</i>	234	20.06	8.47	4.60	47.30

Table 3. Synthetic Table of Social variables

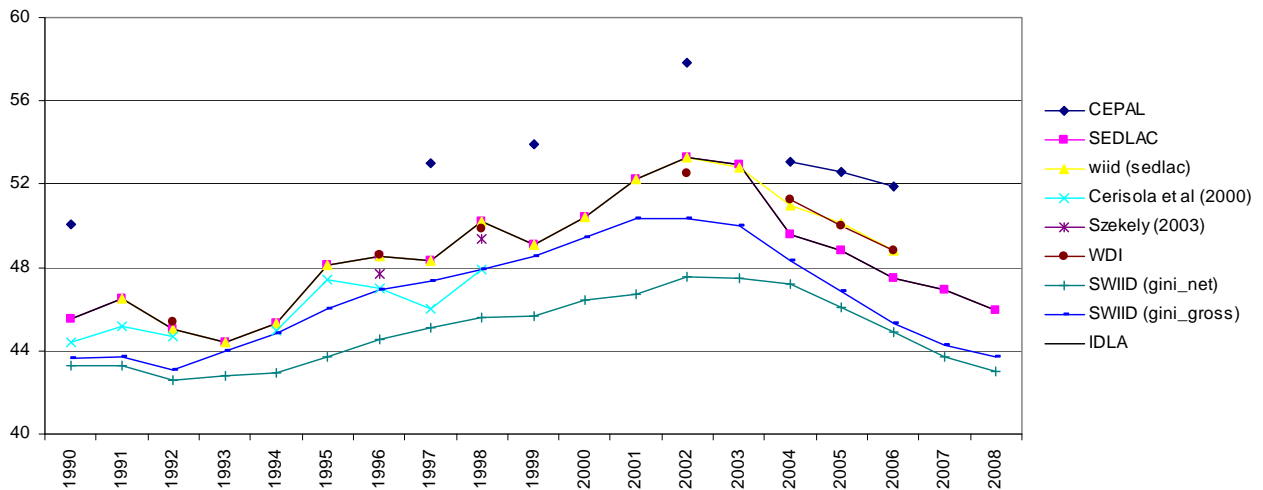
Variable	Obs	Mean	Std. Dev.	Min	Max
<i>gini</i>	323	52.26	4.64	41.20	61.70
<i>pov1</i>	142	3.73	3.42	0.50	20.26
<i>pov2</i>	142	7.79	5.38	0.50	32.64
<i>pov3</i>	142	10.08	7.02	2.00	43.50
<i>pov4</i>	142	19.56	10.25	2.00	61.59
<i>pop</i>	342	26600000	41500000	2412554	192000000
<i>d_pop</i>	342	57.95	67.95	6.16	296.04
<i>urban</i>	72	66.58	15.17	40.30	92.30
<i>life</i>	338	71.22	3.84	58.78	78.92
<i>umr</i>	126	32.09	18.95	8.70	122.10
<i>mmr</i>	90	116.48	77.62	25.00	510.00
<i>fertility</i>	334	3.11	0.83	1.88	5.58
<i>illy</i>	72	13.70	9.88	2.00	39.00
<i>hc_low</i>	202	58.18	12.39	31.90	82.90
<i>hc_medium</i>	202	27.98	7.81	11.50	47.90
<i>hc_high</i>	202	13.84	5.30	3.60	27.90
<i>edu_years</i>	202	7.62	1.54	3.90	10.90
<i>gini_edu</i>	201	0.36	0.09	0.21	0.62
<i>enr_i</i>	200	94.39	4.91	74.10	99.60
<i>enr_ii</i>	200	55.98	17.72	17.00	86.50
<i>enr_iii</i>	200	15.58	7.02	2.40	35.90
<i>mobile</i>	342	18.67	27.01	0.00	116.63
<i>internet</i>	342	6.03	8.84	0.00	40.00
<i>informal</i>	314	51.89	11.89	30.90	77.12
<i>mw</i>	342	102.19	23.67	28.26	252.50
<i>un</i>	292	8.88	4.26	1.30	20.70
<i>lab</i>	342	65.99	5.51	49.35	80.63
<i>wlab</i>	342	36.20	4.11	27.41	43.89
<i>al_ethnic</i>	342	0.43	0.19	0.17	0.74
<i>al_language</i>	323	0.17	0.17	0.02	0.60
<i>al_religion</i>	342	0.28	0.12	0.14	0.61

APPENDIX 2: A comparison between different data of income inequality

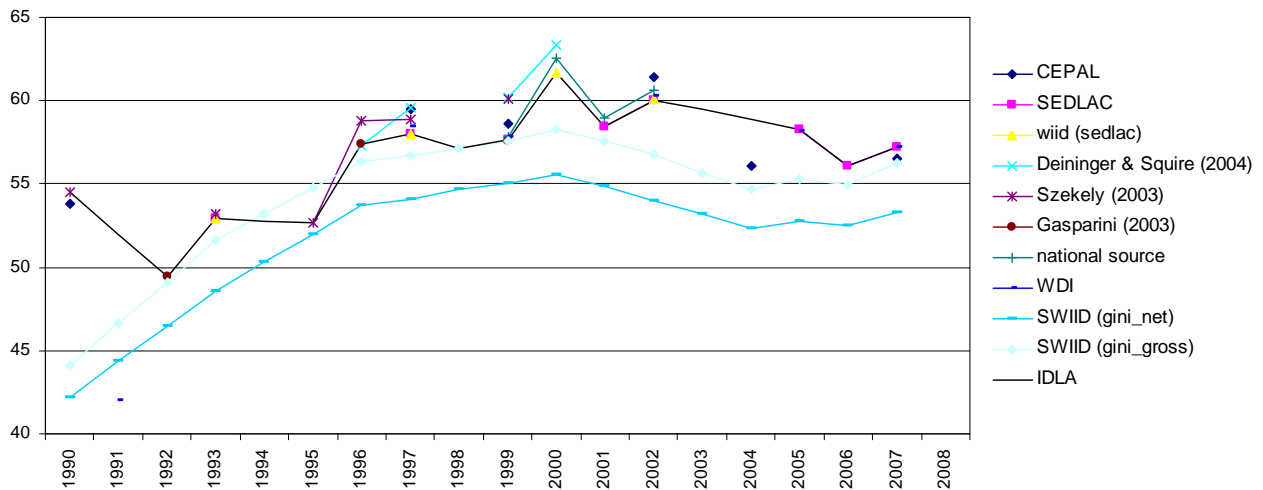
Table 4. Sources of the Gini coefficients included in IDLA

	SEDLAC	WIID				CEPALS TAT	WDI	SWIID	Interpol.	missing	tot
		Szekely (2003)	Gasparini (2003)	SEDLAC (2006)	Szekely & Hilgert (2002)						
Argentina	19	0	0	0	0	0	0	0	0	0	19
Bolivia	7	2	2	1	0	0	0	2	4	1	19
Brazil	15	0	0	0	0	0	0	0	4	0	19
Chile	9	0	0	0	0	0	0	0	10	0	19
Colombia	7	0	0	0	0	1	1	5	5	0	19
Costa Rica	19	0	0	0	0	0	0	0	0	0	19
Ecuador	10	0	0	0	0	0	0	0	4	5	19
El Salvador	15	0	0	0	0	0	0	0	4	0	19
Guatemala	5	0	0	0	0	1	0	8	3	2	19
Honduras	15	0	0	0	0	0	0	2	2	0	19
Mexico	11	0	0	0	0	0	0	0	8	0	19
Nicaragua	4	0	0	0	0	0	0	8	2	5	19
Panama	12	0	0	0	0	0	0	0	7	0	19
Paraguay	11	0	0	0	0	0	0	0	3	5	19
Peru	12	0	0	0	2	0	0	0	4	1	19
Dom. Rep.	11	0	0	0	0	0	0	6	2	0	19
Uruguay	15	0	0	0	0	0	0	0	4	0	19
Venezuela	12	0	0	0	0	2	0	0	5	0	19
Tot	209	2	2	1	2	4	1	31	71	19	342
%	61%	1%	1%	0%	1%	1%	0%	9%	21%	6%	100%

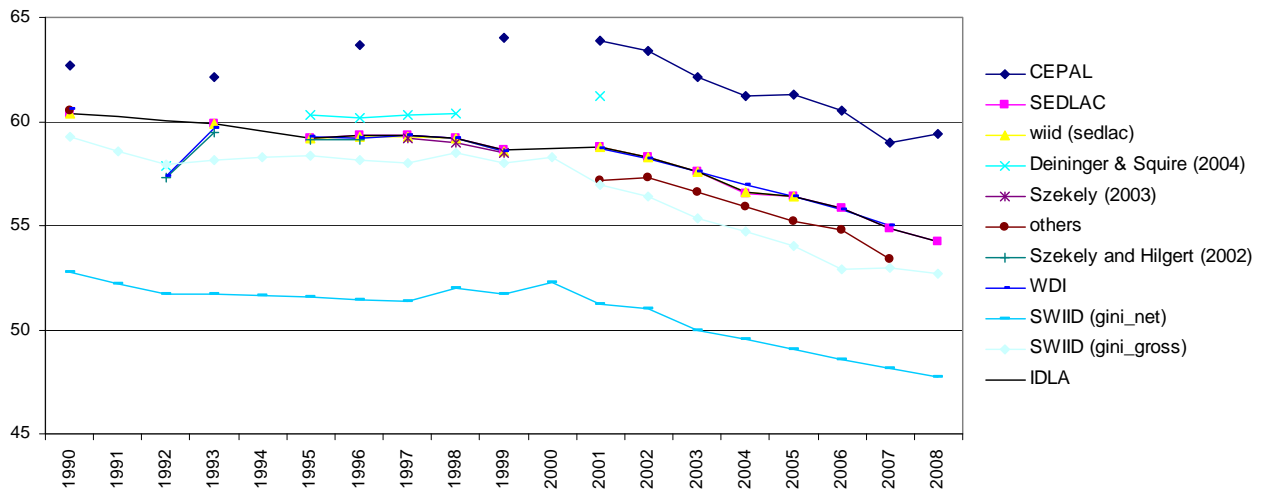
ARGENTINA



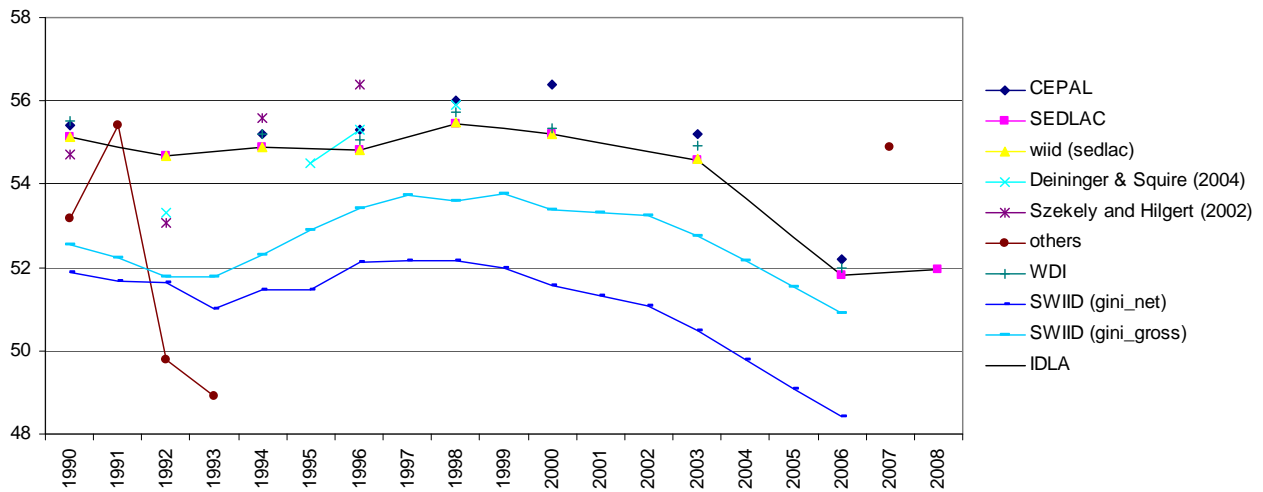
BOLIVIA



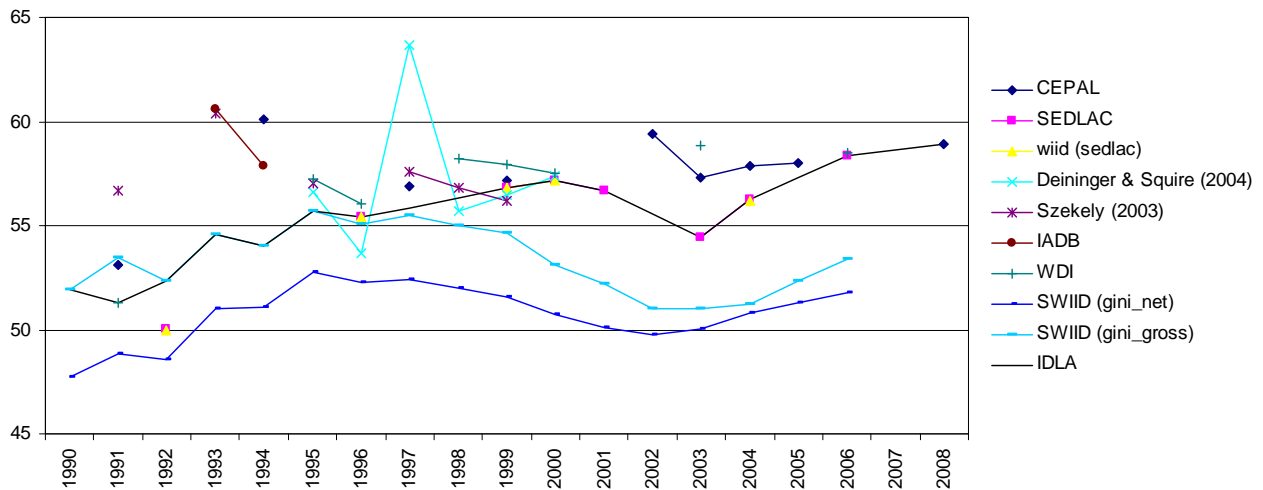
BRAZIL



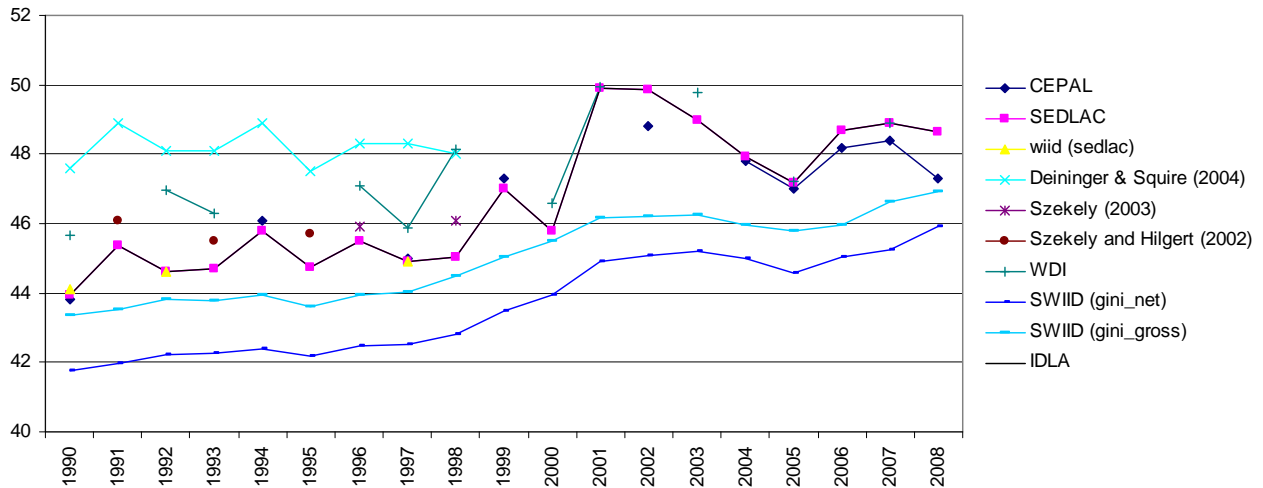
CHILE



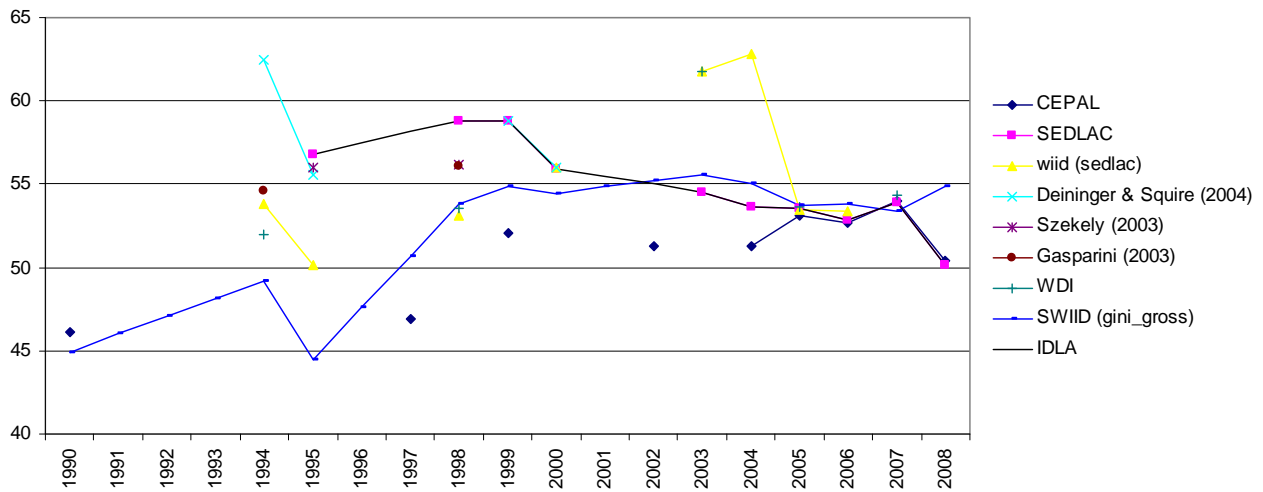
COLOMBIA



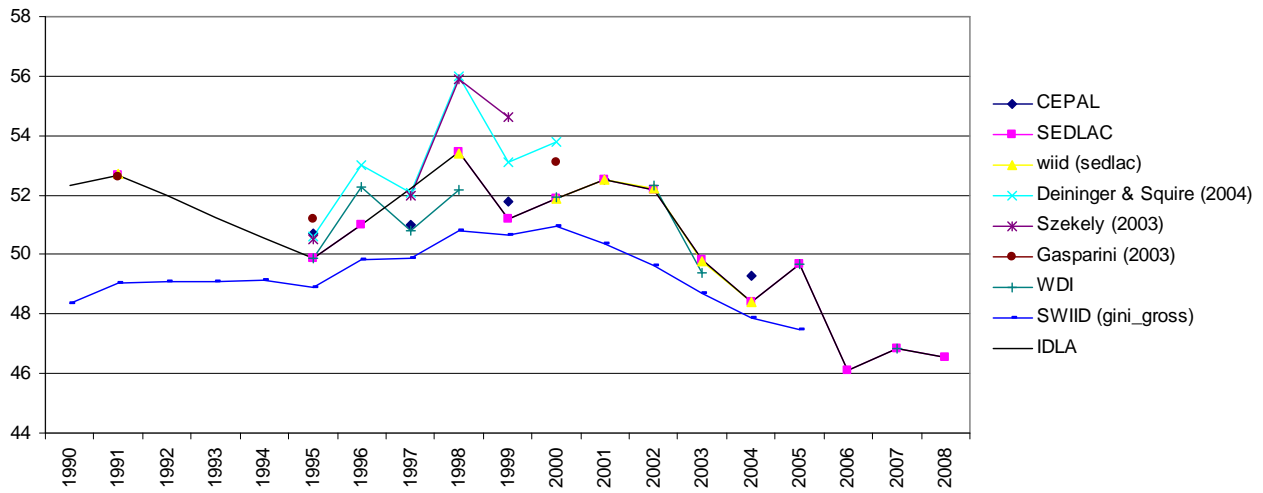
COSTA RICA



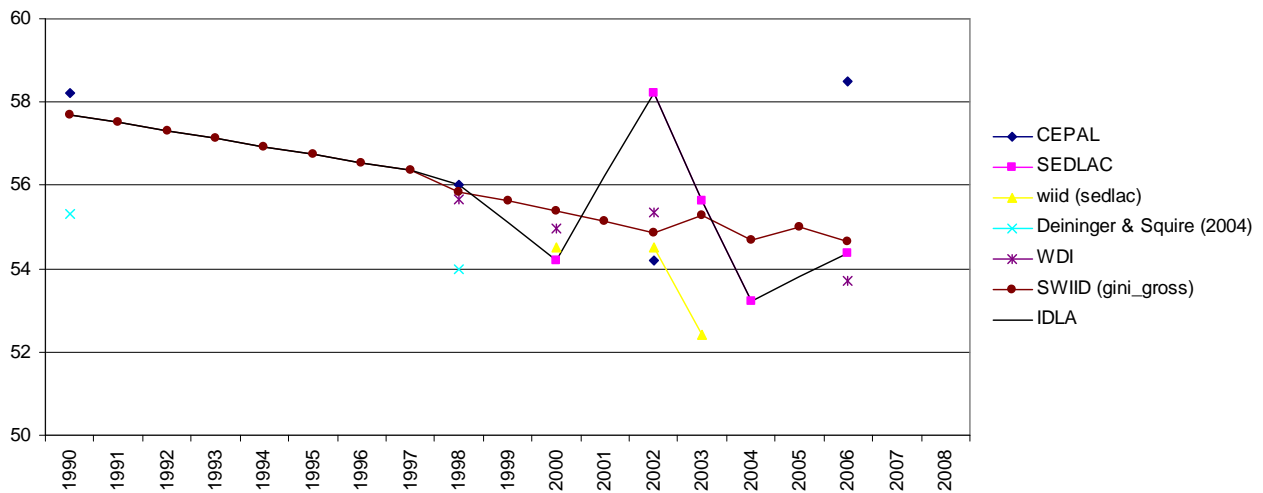
ECUADOR



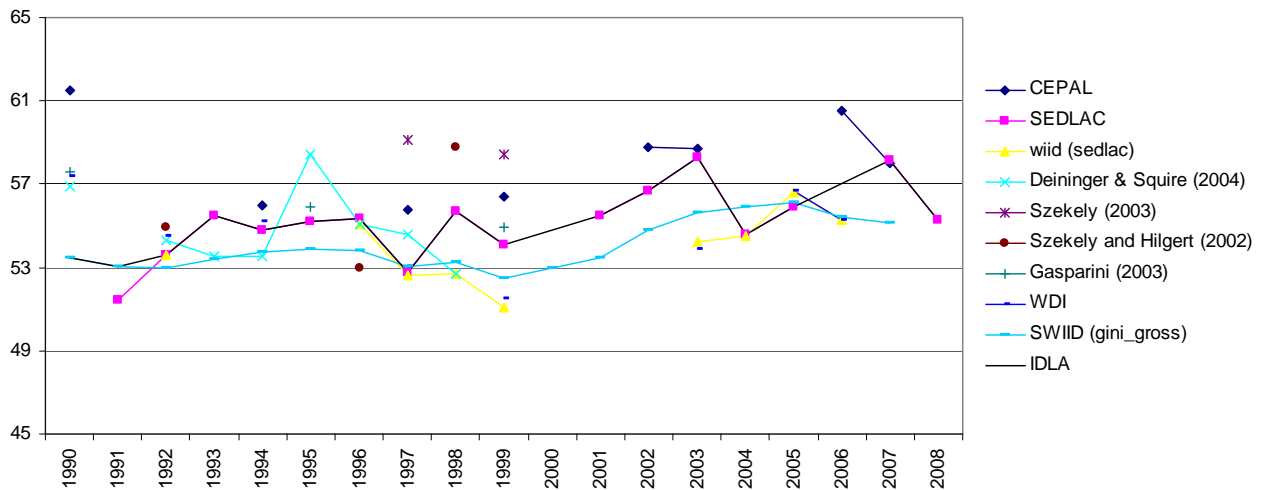
EL SALVADOR



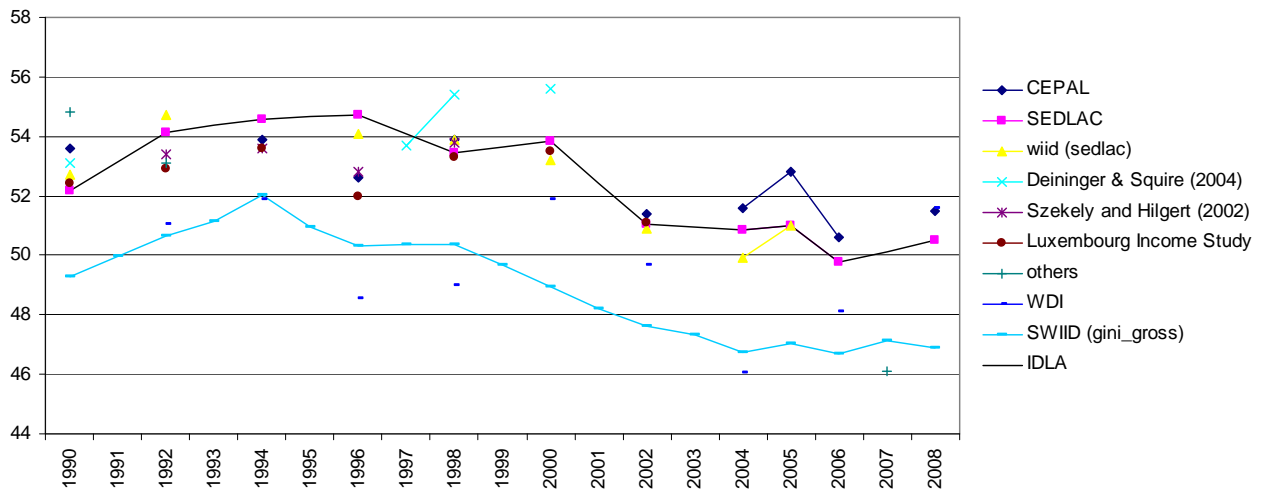
GUATEMALA



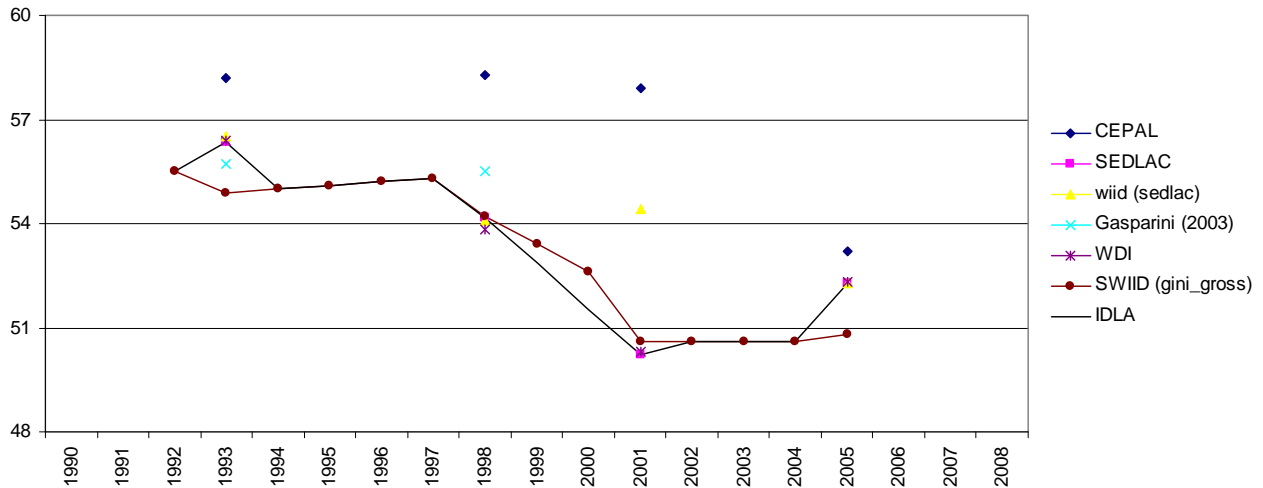
HONDURAS



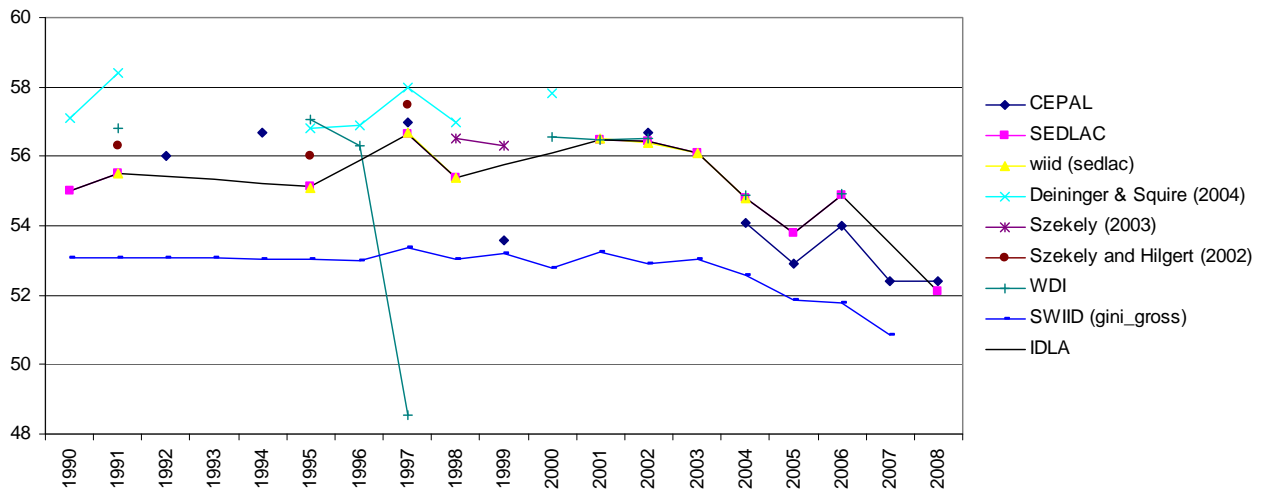
MEXICO



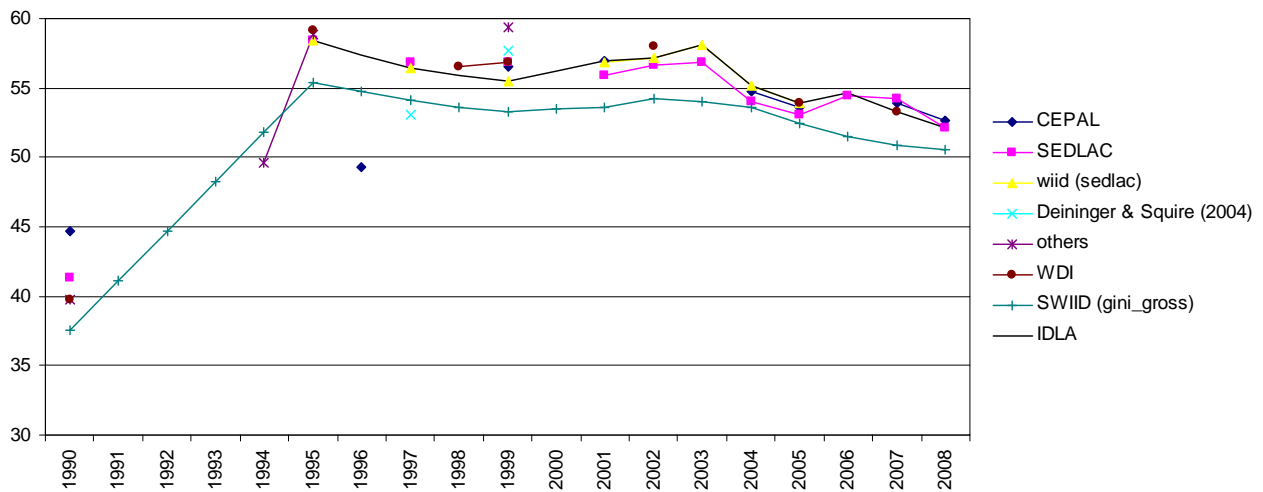
NICARAGUA



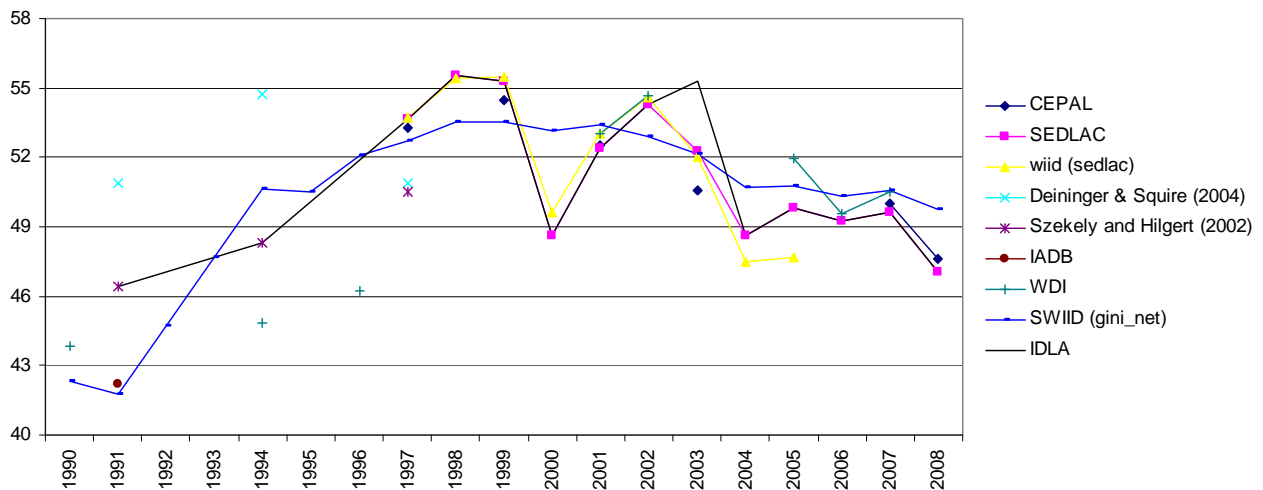
PANAMA



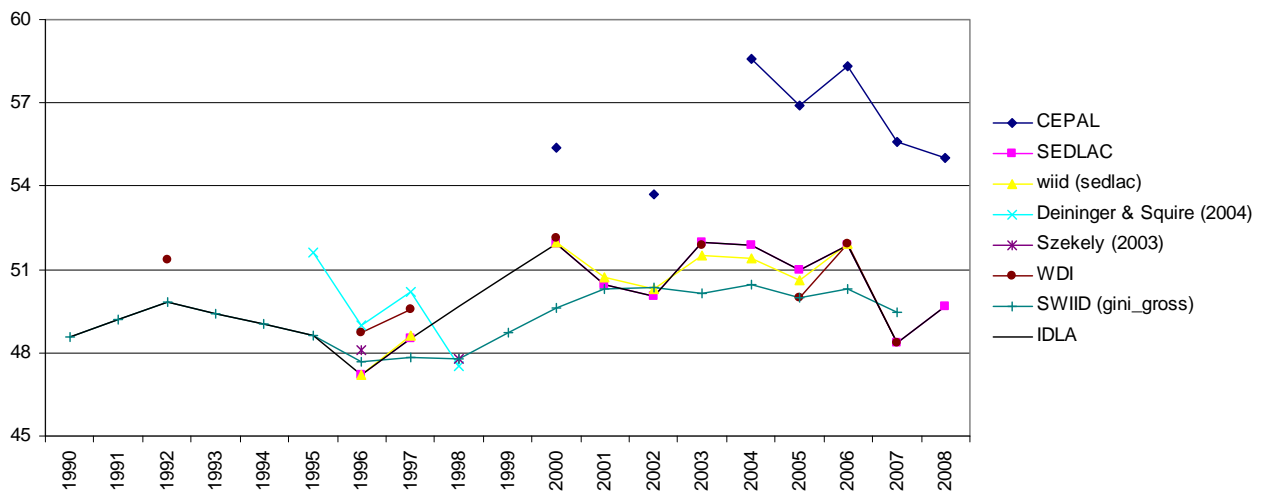
PARAGUAY



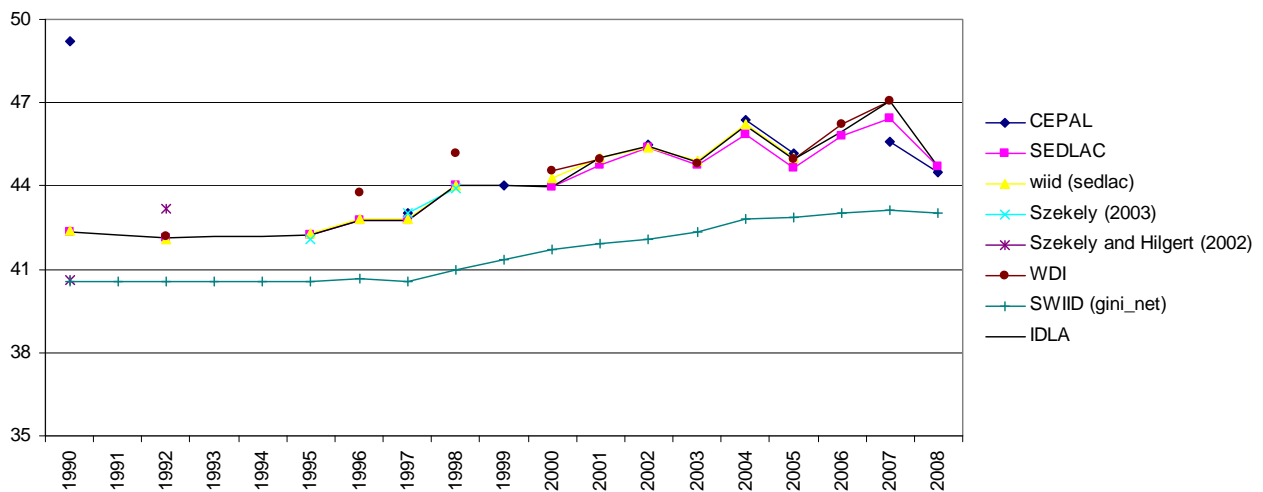
PERU



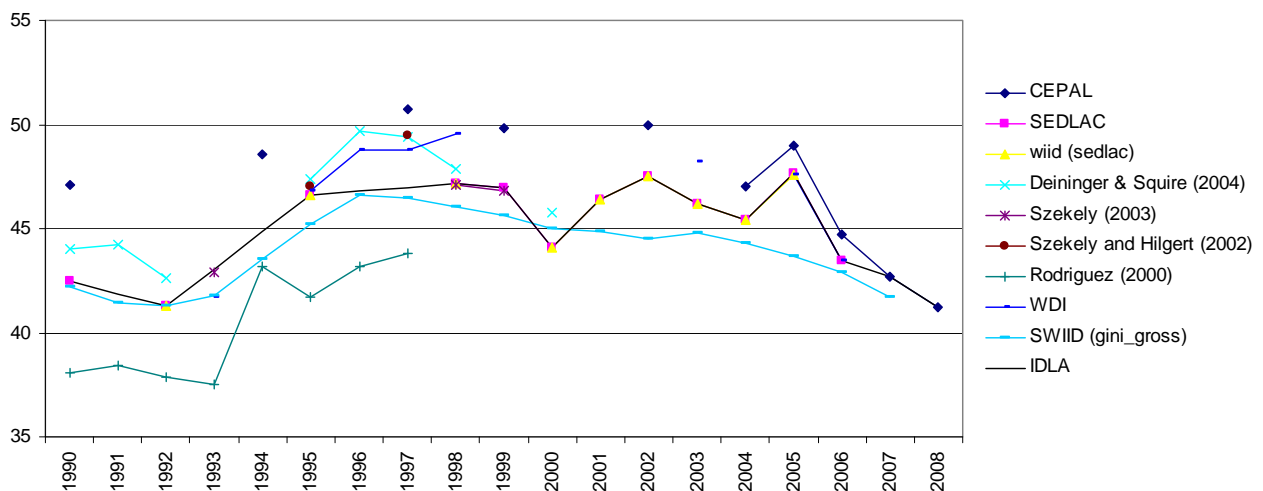
DOMINICAN REPUBLIC



URUGUAY



VENEZUELA



APPENDIX C: Primary Data Sources

Beck, T., Clarke, G., Groff, A., Keefer, P. and P. Walsh (2001), “[New tools in comparative political economy: The Database of Political Institutions](#)”, *World Bank Economic Review*, 15 (1), 165-176.

CEDLAS and World Bank, SEDLAC (Socio Economic Database for Latin America and the Caribbean)
www.depeco.econo.unlp.edu.ar/sedlac/eng/statistics.php

ECLAC, CEPALSTAT database
www.eclac.cl/estadisticas/bases/

ECLAC, Economic Survey of Latin America and the Caribbean, several issues
http://www.eclac.org/cgi-bin/getProd.asp?xml=/de/agrupadores_xml/aes252.xml&xsl=/agrupadores_xml/agrupa_listado.xsl

Frederick, S. (2008-09), “The Standardized World Income Inequality Database”, SWIID Version 3.0
<http://dvn.iq.harvard.edu/dvn/dv/fsolt/faces/study/StudyPage.xhtml?studyId=36908>

ILO, Labour Overview, several issues
http://intranet.oit.org.pe/index.php?option=com_content&task=blogcategory&id=40&Itemid=1155

IMF, WEO (World Economic Outlook database),
<http://www.imf.org/external/pubs/ft/weo/2010/01/weodata/index.aspx>

Kaufmann, D., Kraay, A. and M. Mastruzzi (2009), The Worldwide Governance Indicators (WGI) project,
<http://info.worldbank.org/governance/wgi/index.asp>

Marshall, M. and K. Jaggers, Polity IV Project
<http://www.systemicpeace.org/polity/polity4.htm>

Reinhart, C., M. and K. S. Rogoff (2002), “Background Material to A Modern History of Exchange Rate Arrangements: A Reinterpretation”, International Monetary Fund, Washington DC.

Teorell, J., Samanni, M., Charron, N., Holmberg, S. and B. Rothstein (2010), “The Quality of Government Dataset, version 27May10. University of Gothenburg: The Quality of Government Institute,
<http://www.qog.pol.gu.se>.

UNCTAD, Statistics Overview
<http://www.unctad.org/Templates/Page.asp?intItemID=1584&lang=1>

USAID, Latin America and the Caribbean: Selected Economic and Social Data
<http://lac.eads.usaidallnet.gov/>

USDA, ERS International Macroeconomic Data Set
<http://www.ers.usda.gov/Data/Macroeconomics/>

Vanhanen, T., Measures of Democracy 1810-2008. FSD1289, version 4.0 (2009-12-14). Tampere: Finnish Social Science Data Archive, 2009.
<http://www.fsd.uta.fi/english/data/catalogue/FSD1289/meF1289e.html>

WIDER, WIID (World Income Inequality Database)
www.wider.unu.edu/wiid/wiid.htm

World Bank, WDI (World Development Indicators)
<http://data.worldbank.org/data-catalog/world-development-indicators>