Inequality and Structural Change

Hyunsub Kum

Seoul National University

Seoul

Republic of Korea

hyun.kum@gmail.com

September 22, 2008

University of Texas Inequality Project

UTIP Working Paper No. 54

Abstract:

This paper presents an updated data set on inequality in structures of manufacturing pay for the years 1963 – 2002, using the standard methods of the University of Texas Inequality Project (http://utip.gov.utexas.edu). The paper then compares these measures with evidence on structural change, taken as changing shares of agriculture, manufacturing and services in total employment. A key finding is that low inequality is closely associated with low variability in inequality through time, and that movement out of agriculture is associated with high variability in the inequality of manufacturing pay. Thus the level of inequality is a reasonable index of underdevelopment, and the change of the UTIP inequality measure is an indicator of overall structural change in the process of development.

1. Introduction

This paper presents an updated data set on inequality in structures of manufacturing pay for the years 1963 – 2002, using the standard methods of the University of Texas Inequality Project (http://utip.gov.utexas.edu). The paper then compares these measures with evidence on structural change, taken as changing shares of agriculture, manufacturing and services in total employment. A key finding is that low inequality is closely associated with low variability in inequality through time, and that movement out of agriculture is associated with high variability in the inequality of manufacturing pay. Thus the level of inequality is a reasonable index of underdevelopment, and the change of the UTIP inequality measure is an indicator of overall structural change in the process of development.

2. Data for the Measurement of Pay Inequality

Data on inequality for this study are derived from the Industrial Statistics Database of United Nations Industrial Development Organization (UNIDO),[1] which provides total payroll and annual average employment according to International Standard Industrial Code (ISIC) Revision 2 at the 3-digit level. This comprises 28 manufacturing industries for 155 countries in the 1963 – 2003 period. From this we compute 3,452 observations on pay inequality in manufacturing industry in somewhat consistent standardized format covering nearly forty years. These data have several merits for comparative analyses in cross-sections and time-series.

---

[1] This study uses the 2005 version of the UNIDO Industrial Statistics data set.

First, the data have been collected and managed in a consistent manner by UNIDO for a long time. All measures of pay and employment -- the necessary ingredients of the UTIP-UNIDO measure of inequality -- have been collected as a matter of official routine by each government following ISIC 3-digit framework in most countries around the world. Pay is defined as "wages and salaries paid to employees in a year" and employment is as "employees" or "persons engaged" by UNIDO criteria. This simplicity may minimize the noise associated with varying interpretations of the definition. Table 1 shows the detail of 3-digit ISIC industry classifications, which is used as the framework for aggregation.

Table 1. Manufacturing Sectors by 3-digit ISIC Code

ISIC	Industry	ISIC	Industry
311	Food production	354	Misc. petroleum/coal production
313	Beverages	355	Rubber production
314	Tobacco	356	Plastic production
321	Textiles	361	Pottery/china/earthenware
322	Wearing apparel, w/o footwear	362	Glass/ production
323	Leather production	369	Other non-metallic mineral production
324	Footwear, w/o rubber or plastic	371	Iron/steel
331	Wood production, w/o furniture	372	Non-ferrous metals
332	Furniture, w/o metal	381	Fabricated metal production
341	Paper/ production	382	Machinery, w/o electrical
342	Printing/ publishing	383	Machinery electric
351	Industrial chemicals	384	Transport equipment
352	Other chemicals	385	Professional/Scientific equipment
353	Petroleum refineries	390	Other manufactured production

Second, all values for pay and employment in this data are measured in annual terms. Of course, the annual *average* pay is a rough measure, which might be affected by changes in the length of work-time, in numbers of part-time workers, or change in the gender composition of the workforce. Also, there are still conceptual differences in annual pay or its calculation among different countries. This is because pay may include not only direct measures of "wages and salaries" but also several "auxiliary benefits paid to employees" (for instance social security, pension, insurance, or severance pay), which are different from country to country.[2]

However, when comparing the annual average pay from the UNIDO data with the average hourly compensation costs from the US. Bureau of Labor Statistics,[3] which are constructed for the assessment of international differences in employer labor costs, the correlation coefficients in the cases of OECD countries are above 0.95 except for France (0.82) and Mexico (0.72).[4] Thus, we can borrow some strength from the ICHCC to check the cross-country comparability of the annual average values in the UNIDO data. Further, the fact that most countries stick to their reporting conventions and statistical procedures over time allows us reasonably to expect the comparability of measures over time within a country. Berman's endorsement (2000) of the coverage and accuracy of the UNIDO compilation lends some weight to our confidence in the quality of this data set.

[2] Pay and salaries in terms of UNIDO's definition include "all payments in cash or in kind made to employees during the reference year in relation to work done for the establishment."

[3] This is the International Comparisons of Hourly Compensation Costs for Production Workers in Manufacturing (ICHCC) data, which provides average labor compensation costs for 28 countries in 1975-2000 at five year intervals. Rodrik (1999) took the same approach to check the quality of UNIDO data.

[4] Countries in this comparison include Australia, Austria, Belgium, Canada, Germany, Denmark, Spain, Finland, France, Greece, Hong Kong, Ireland, Israel, Italy, Japan, Korea, Sri Lanka, Luxembourg, Mexico, Netherlands, Norway, New Zealand, Portugal, Singapore, Taiwan, Sweden, the United Kingdom and the United States.

3. Theil's T Inequality Measure

The UTIP-UNIDO measure of inequality is the between-groups component of a Theil generalized entropy index of inequality, which has perfect decomposability into between-group (T_B) and within-group (T_w) components as shown below.[5] If we divide our subject pool into several groups, T_w is a weighted average of the Theil index for each group, and T_B is a weighted geometric mean of the wage relativities, using the share of aggregate pay as a weight.

<u>Theil Inequality Index and its Decomposition</u>

$$T_{total} \equiv T_B + \overline{T}_W$$

$$T_B = \sum_{i=1}^{n} \frac{w_i}{\sum_{i=1}^{n} w_i} \bullet \ln \left[\frac{w_i \Big/ \sum_{i=1}^{n} w_i}{e_i \Big/ \sum_{i=1}^{n} e_i} \right]$$

$$T_w = \left(\frac{w_{ij}}{w_i} \right) \bullet \left[\frac{w_{ij} / w_i}{e_{ij} / e_i} \right]$$

$$\overline{T}_w = \sum_{i=1}^{n} \frac{w_i}{w} \bullet T_w$$

We focus on the group-wise inequality or between-group inequality (T_B) component, which requires only group-wise measures (means of pay and employment) without any further information. With these, the calculation of the measure of inequality is straightforward as shown in the above formula. Also since this measure is a distance function showing divergence between wage shares and employment shares by groups, the changes of pay and employment are explicitly reflected in the calculations of change

---

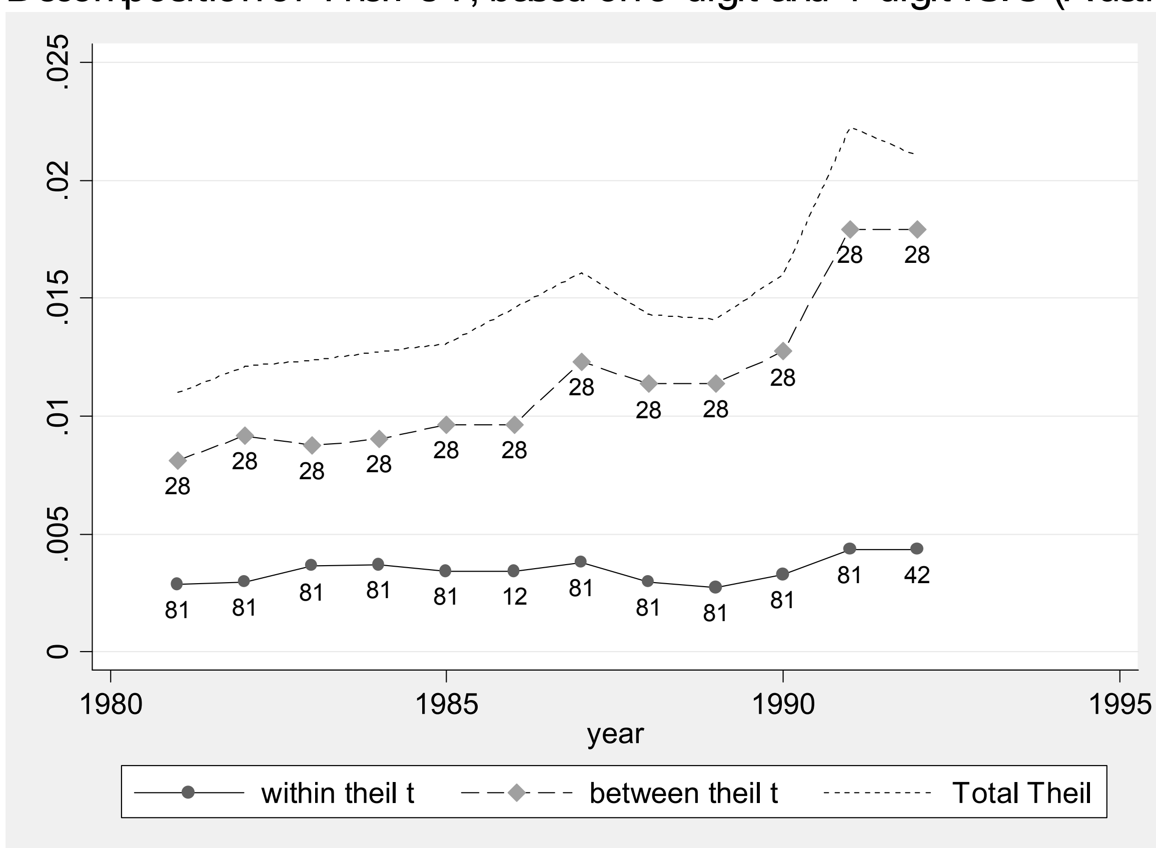
5 The popular Gini inequality index also can be decomposed into between, within, and overlap components (Pyatt, 1976). However, in this case, the overlap component cannot be identified from aggregated measures alone, thus only an approximation of the between-groups component is available.

over time. The underlying grouping scheme can be just about anything -- gender, race, economic sector, or geographic region – so long as the groups are mutually exclusive and collectively exhaustive (MECE). In the UNIDO data, 3-digit ISIC (International Standard Industrial Classification) code for manufacturing industry meets this specification, and has the added virtue of placing wage and employment changes that reflect structural change in the economy into the between-group component of inequality where it can be directly observed; with other classification schemes, such as gender or region, it is possible that structural change would be reflected mainly in the within-group element of inequality, which is unobserved.

One may still ask whether omitting the within-industry component would make a significant difference to our understanding of the underlying economic processes. Without doubt, the degree of approximation of T_B to T_{total} may depend on the size of the within-industry component for each country and year. But Theil (1972) argued that an inequality measure computed from grouped data provides a *consistent lower-bound* estimate of inequality for the total population. And a series of empirical studies (Conceicao, Galbraith and Bradford, 2001) shows that T_B is usually a good estimate of changes in the whole distribution when industrial sectorization is employed. Thus, it seems reasonable to assume that the movement of the between-industries component of Theil's T (T_B) approximates the movement of total inequality, especially for the secular trend rather than the absolute level.

To see this point clearly, we combine 3-digit and 4-digit industries data into a hierarchical structure and treat them as between-group (3-digit ISIC) and within-group (4-digit ISIC) components of a common classification. We then calculate the Theil index with the two components. Figure 1 shows the Australian case. When the Theil index is decomposed into T_B and T_W in this way, the relative magnitude of the latter is much smaller than that of the former, and the former very sufficiently represents the overall trend (Data labels show the number of categories available in each year.)

Figure 1. Decomposition of Theil's T, based on 3-digit and 4-digit ISIC (Australia)



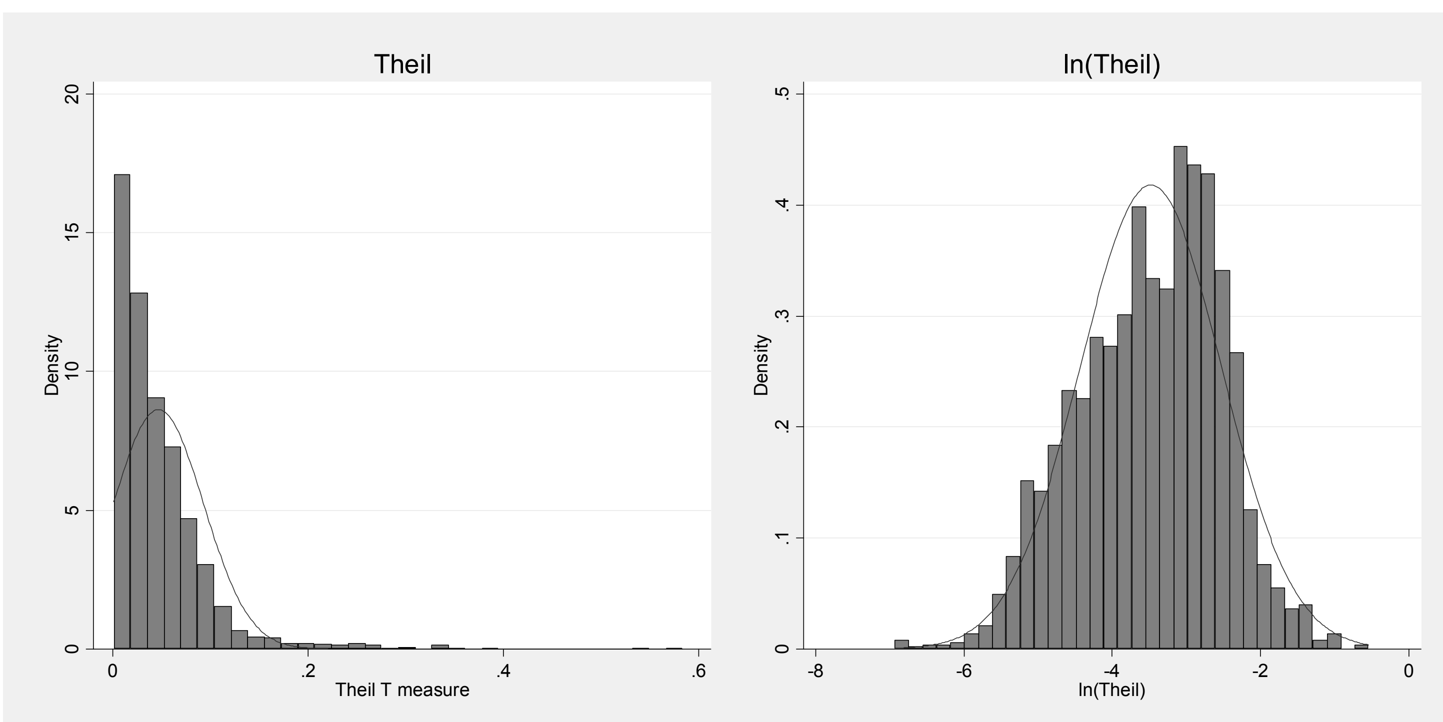
Further disaggregation, carried out for the US by Conceicao, Galbraith and Bradford (2001) confirms that moving to finer levels of disaggregation yields diminishing returns in information about the movement of inequality: the fine classification schemes tend to have the same broad features as the coarse schemes, just as a low-resolution photograph captures the broad features of a landscape while a high-resolution picture merely adds detail. Thus it can be said that changes in the between-industry component do arguably provide a useful approximation of the changes in overall industrial pay inequality in the majority of countries and time periods covered in this study.

Based on the 2005 release of UNIDO's ISIC, we calculate 3,452 Theil T inequality measures for 155 countries within the 1963-2003 period. The distribution of our measure across regions and time by decade are tabulated in the Table 2, and Figure 2 presents information on the distribution of the data in raw and logged form. It is interesting that the distribution is approximately log-normal; alternatively it may be said to resemble the statistically most-probable Boltzmann distribution.

Table 2. Distribution of UTIP-UNIDO Inequality Measures

Year	East Asia & Pacific	East & Central Asia	Latin & Central America	Middle East & North Africa	North America	South Asia	Sub-Saharan Africa	Western Europe	Total
Country	19	24	28	7	2	8	38	19	145
1960s	49	52	75	62	14	25	83	113	473
1970s	118	80	162	102	20	46	205	183	916
1980s	142	93	192	96	20	64	198	190	995
1990s	125	161	162	97	19	43	147	173	927
2000s	21	40	10	20	4	9	18	19	141
Total	455	426	601	377	77	187	651	678	3,452

Figure 2. Distributions of UTIP-UNIDO Inequality Measures



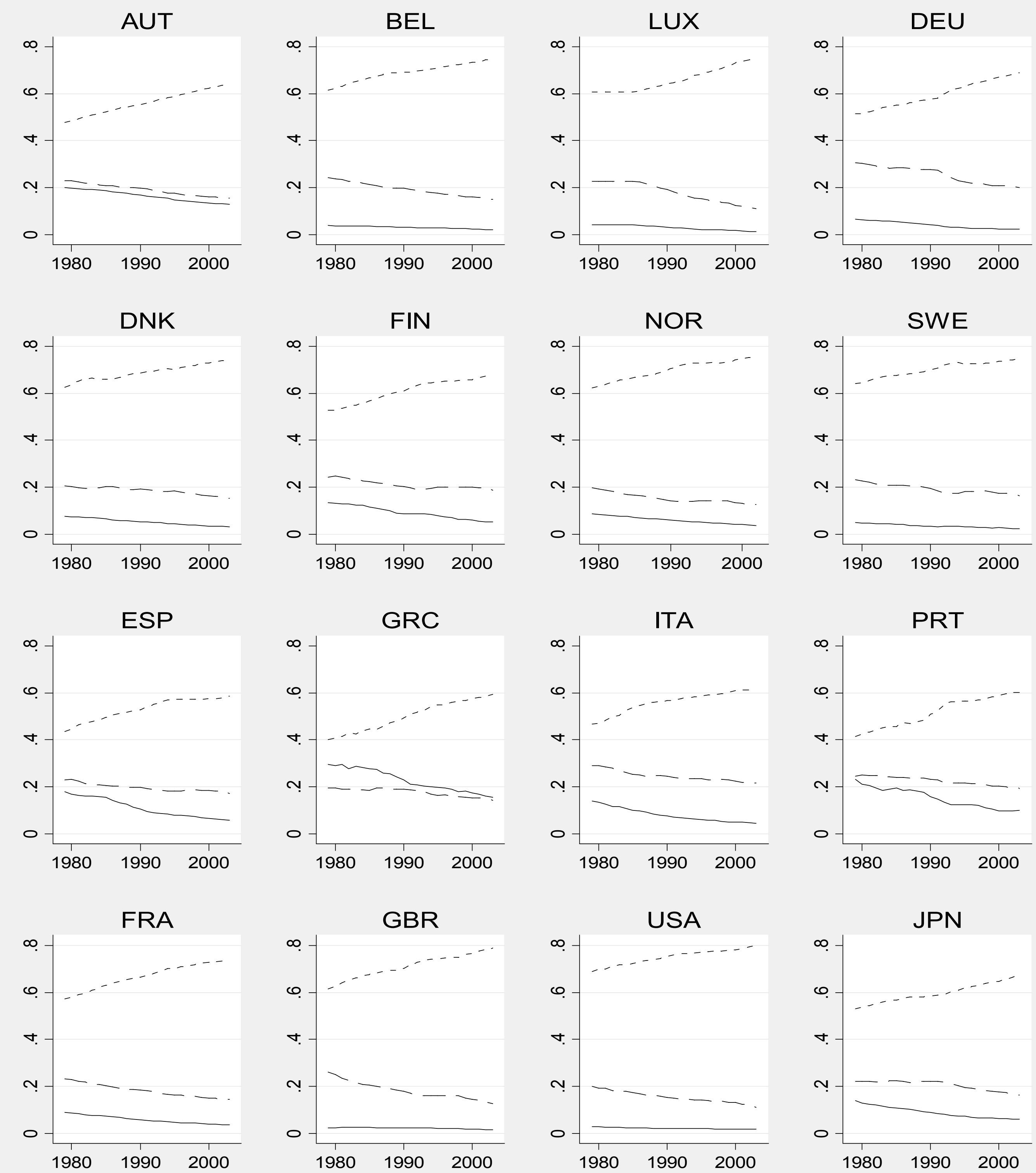
4. Employment Shares and Inequality

The sectoral employment shares for agriculture, manufacturing and services for selected countries are presented below. Kuznets (1955) identified the transition from agriculture to industry as the main source of change of increasing inequality in the early stage of economic development. As the size of agriculture shrinks, the size of industry grows, and inequality increases due to the large gap between the two sectors. But as the agricultural sector falls to an unimportant share of total employment, then trends inside the industrial sector come to dominate, and with income growth and the development of social democracy inequality takes a decreasing path; thus emerges the Kuznets inverted curve relating inequality to income. Clearly inter-sectoral migration of labor is one notable factor in the evolution of inequality, and the change of employment share by sectors or inter-sectoral transitions in employment could shed light on the underlying change of economic structure.

Figure 3 presents employment shares for agriculture, manufacturing and services for the OECD countries in 1979-2003.[6] It is apparent that the employment share of agriculture has fallen in every country and every period of time, and even more in Spain, Greece, and Portugal, which have been known as agrarian European countries. But it is not clear that workers leaving agriculture move to the manufacturing industries directly. It seems that some would, but others migrate to the services sector, whose increase in employment share is prominent in most OECD countries.

[6] The comparisons of employment share by sectors are based on the data from the Groningen Growth and Development Center (GGDC). Especially we extract the 10 sector database for the sectoral employment share for 42 countries in 1979-2003 for this comparison. See the website http://www.ggdc.net/

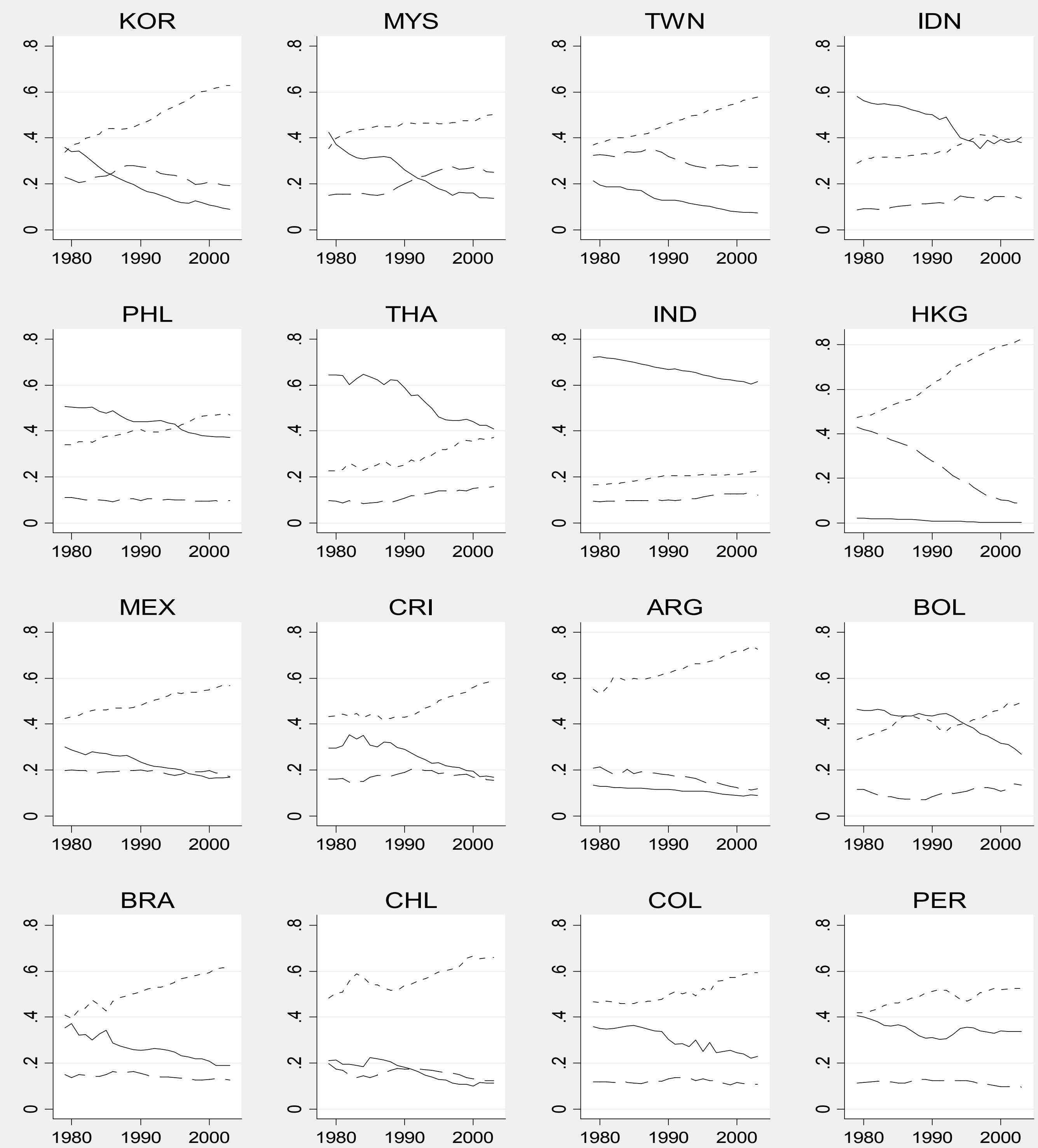
Figure 3. Employment Share by Agriculture, Manufacturing, and Services for Selected OECD Countries



Source: Groningen Growth and Development Center
*Solid for Agriculture, Dashed for Manufacturing, Thin-dahsed for Service

The next figure for selected Asian and Latin American countries indicates a somewhat different pattern. In Asian countries, Korea experienced rapid growth in employment in the services sector, but shrinkage in agriculture. Malaysia and Taiwan show similar changes. The same is true for Indonesia, Philippine and Thailand, but only in later years. India is a major exception; there the employment share of agriculture decreases but at a very slow rate, while growth of other sectors is also very slow. Hong Kong is another exception. In that case, the shrinking sector is not agriculture but manufacturing. Korea and Taiwan also show declines in manufacturing in the later 1990s. In other Asian countries, however, the manufacturing sector shows steady growth.

Figure 4. Employment Share by Agriculture, Manufacturing, and Services for Asian and Latin American countries

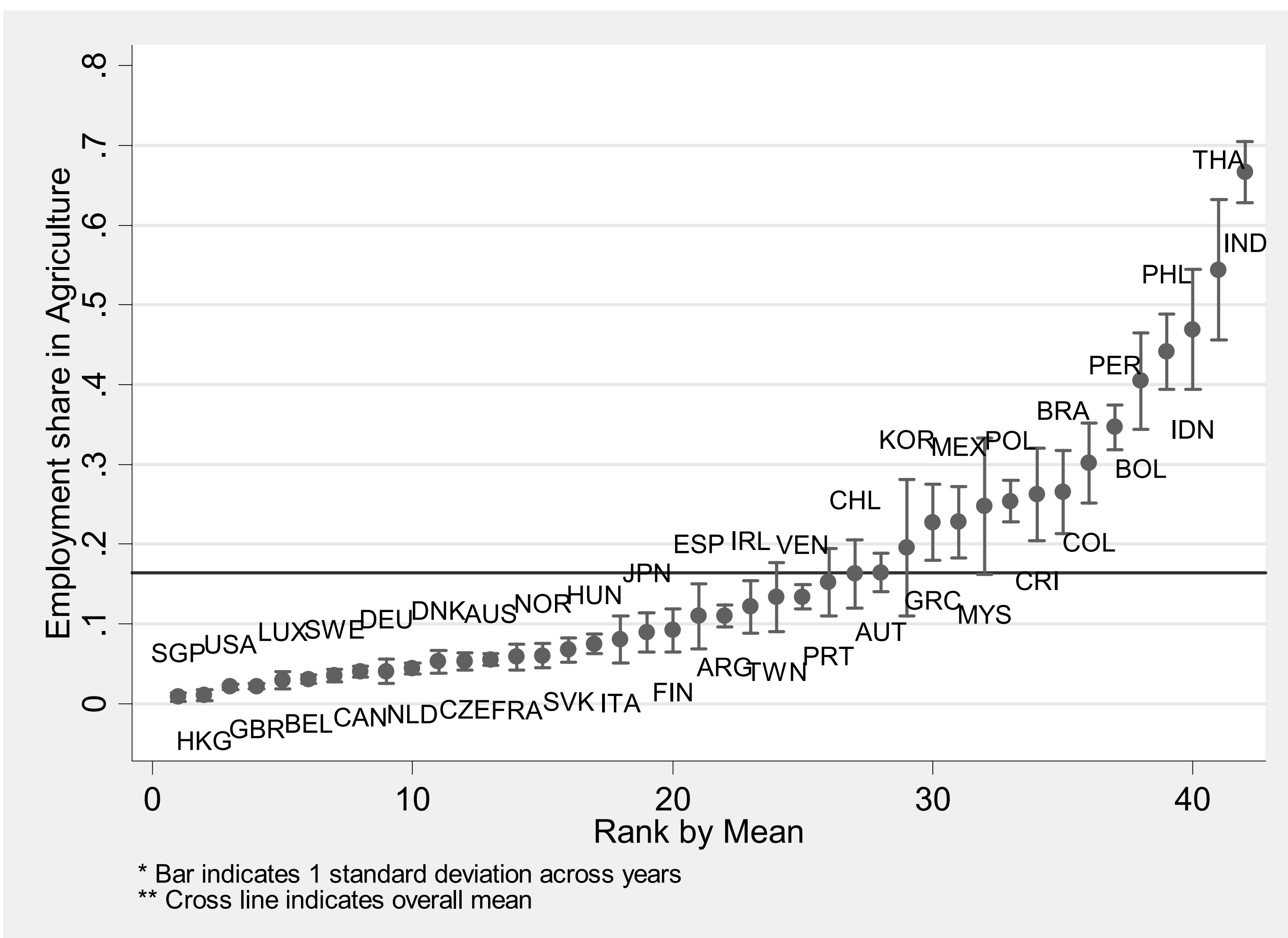


Source: Groningen Growth and Development Center
*Solid for Agriculture, Dashed for Manufacturing, Thin-dahsed for Service

As Figure 4 shows, in most Latin American countries the manufacturing sector employment share has been stagnant for a long time. Mexico and Bolivia have small increases in manufacturing employment share and Argentina shows a decrease. For the service sector, an increasing pattern is prevalent but with some fluctuations that is different from other regions showing monotonic increase. But the agriculture sector has lost a huge amount of labor, in relative terms, in the last 25 years.

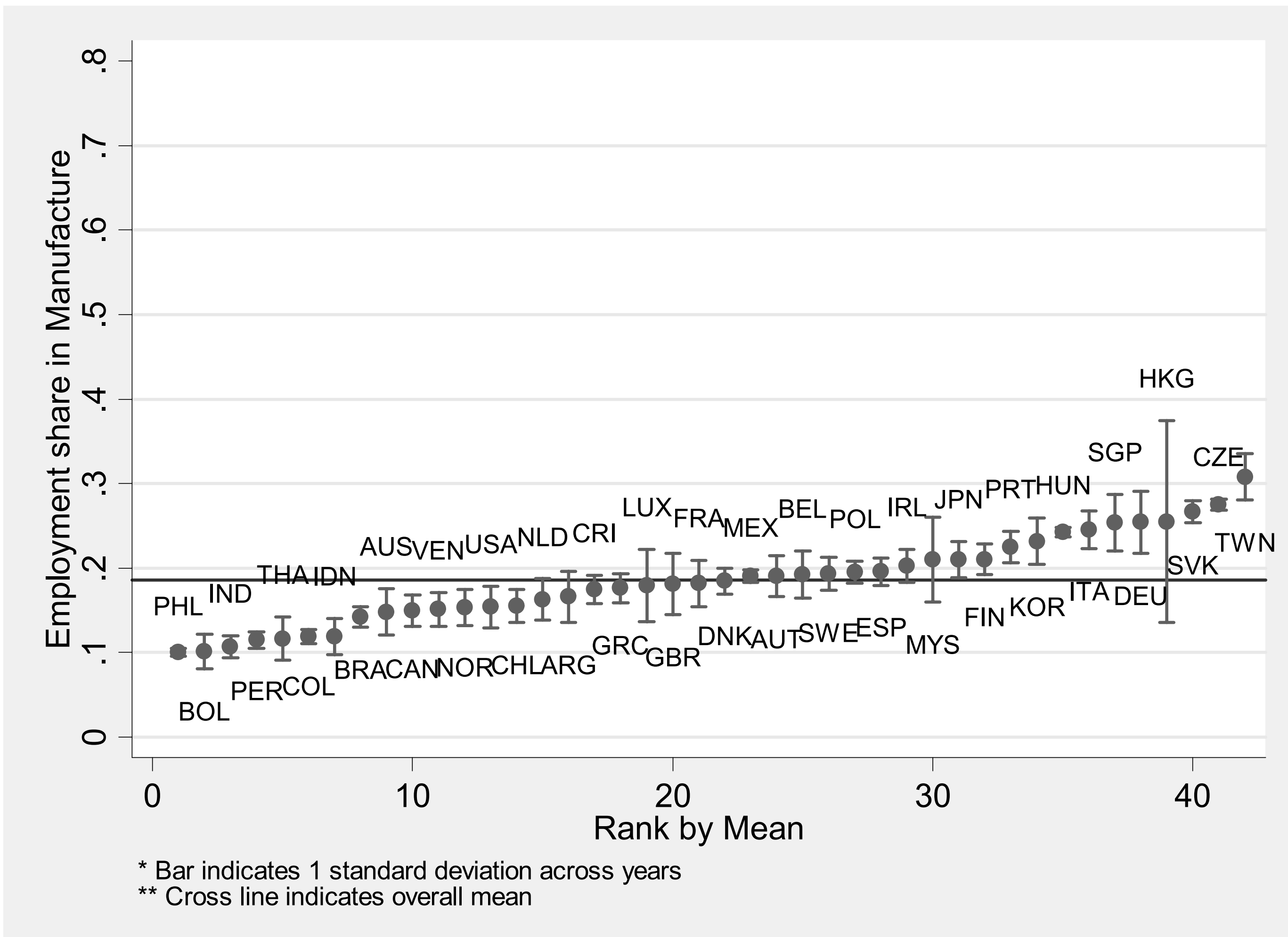
To see the magnitude of change in comparative way, we select 42 countries that have non-missing information for a 10-sector classification for 1979-2003. These include 25 OECD countries and 17 developing countries. Then, we calculate the mean and standard deviation of the employment share by sector for each country as a measure of variation over time within each country. We then arrange them in the order of mean value of employment share for agriculture, manufacturing, and services as shown in Figures 5–7.

Figure 5. Employment Share of Agriculture, 1979-2003



The mean employment share of agriculture in the 42 countries is less than 20%, and countries with smaller shares experience smaller changes. In contrast, countries with higher employment share of agriculture tend to have larger reductions in the agricultural labor force. For instance, Indonesia has a 46.9% employment share in agriculture on average but its standard deviation is 7.5%, whereas the UK has 2.2% of agriculture employment with only 0.3% standard deviation in last 25 years. Thailand, Korea, and Malaysia lost the largest share of the labor force in agriculture in this period.

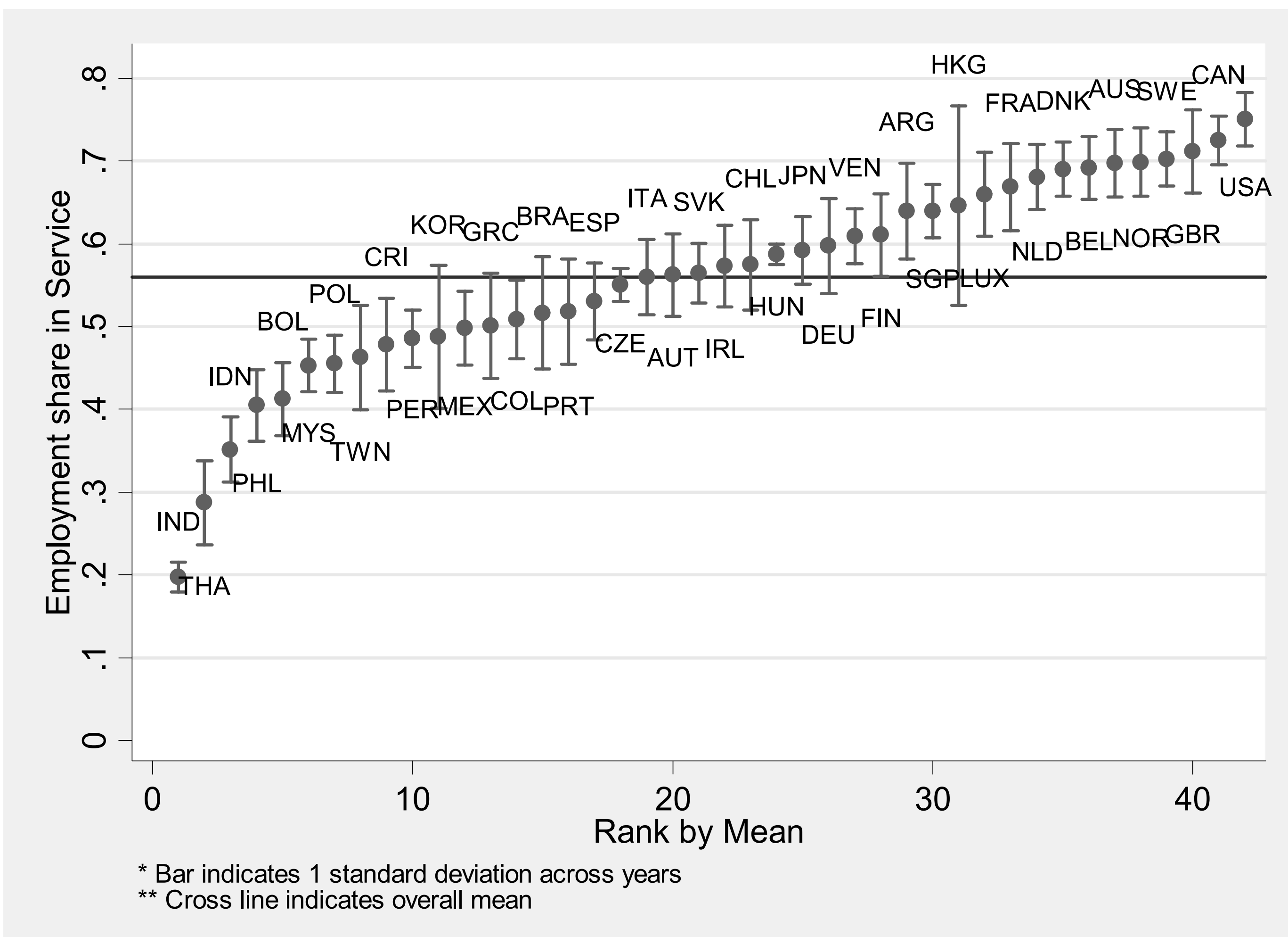
Figure 6. Employment Share of Manufacturing, 1979-2003



When it comes to the manufacturing sector (Figure 6), the employment share is around 20% on average and the estimates of standard deviation are relatively small. Most countries have at least 10% but less than 30% of employees in the manufacturing sector, and only handful of countries such as Hong Kong (11.9%), Malaysia (5%) experience relatively large changes in employment share in this period; the former decreasing and the latter increasing. At this stage, it is hard to find a direct linkage between labor migration out of agriculture and into manufacturing sectors.

Figure 7 is for the services sector. The employment share of the services sector in most countries has been growing rapidly. Financially-advanced countries such as the USA, Canada, and UK (GBR) all take the highest rank, and keep increasing the share of this sector in total employment. But most developing countries experience more volatile change in the share of services in employment. Two exceptions are noteworthy. Hong Kong, a small economy but dominated by finance and real-estate, increases the employment share from 47% in 1979 to 83% in 2003. In Thailand, hard hit by the Asian financial crisis in 1997, services have grown to be just equal to the share of agricultural employment in 2005 (see also Figure 4).

Figure 7. Employment Share of Services, 1979-2003



When the same exercise is made on to the UTIP-UNIDO inequality measure, another interesting finding emerges. We calculate the mean and standard deviation of the inequality measure and place them in the order of their rank based on mean value. For this exercise, we restrict ourselves to countries that have at least 20 annual observations in the 1963-2002 period, resulting in 88 countries. From this, we select the 42 countries that have employment share by sector, as discussed above, for comparison.

Figure 8. Mean and Standard Deviation of UTIP-UNIDO Inequality, 1963-2002.

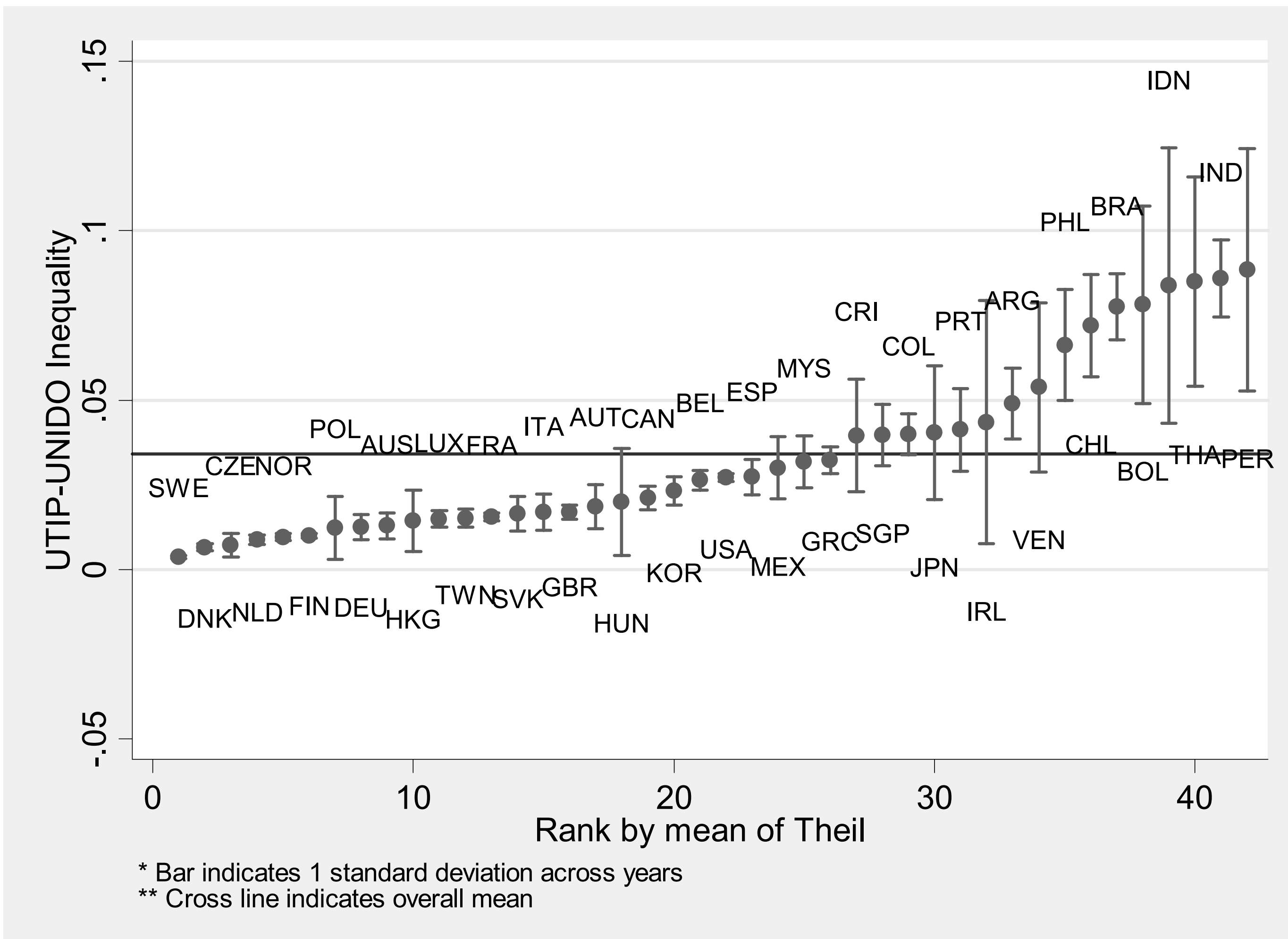


Figure 8 shows the result. The first thing we can notice is that the higher the inequality level is, the larger its standard deviation over time. For instance, several northern European countries have lowest inequality level, as well as the smallest fluctuations in this set of countries over the past 40 years. In contrast, several Latin American countries have much higher levels of inequality with much higher volatility in its change. The Pearson correlation coefficient between average and standard deviation across countries is over 0.77.

This close relationship between the inequality level and its fluctuation is reassessed in Figure 9. Here we plot the rank order based on standard deviation against the rank order based on mean of inequality. Again the relationship clearly emerges: lower ranked countries in inequality level have lower rank in its change over time, whereas higher ranked countries in inequality level have higher rank in its change.

Figure 9. Ranks of Average Level and Standard Deviation in Inequality, 1963-2002.

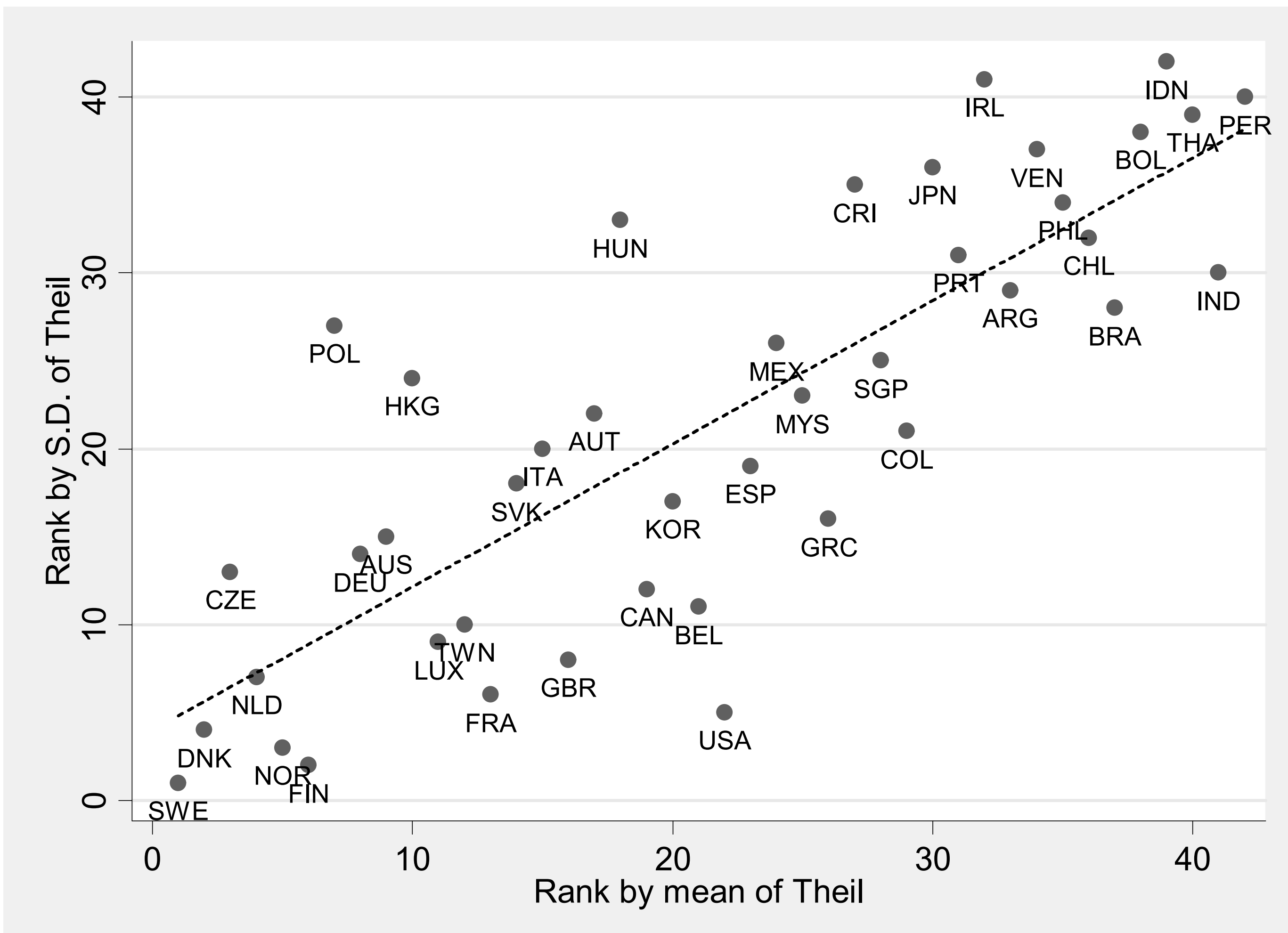


Figure 10 combines the change of inequality and the change of employment share of agriculture for 42 countries, to examine the relationship between those two changes. If some countries that have greater change in employment share in agriculture experience greater change in inequality, then it could indicate some connection between inter-sectoral transition and inequality change.

Figure 10. The Change of Employment Share in Agriculture and the Change of UTIP-UNIDO Theil Inequality, 1979-2002.

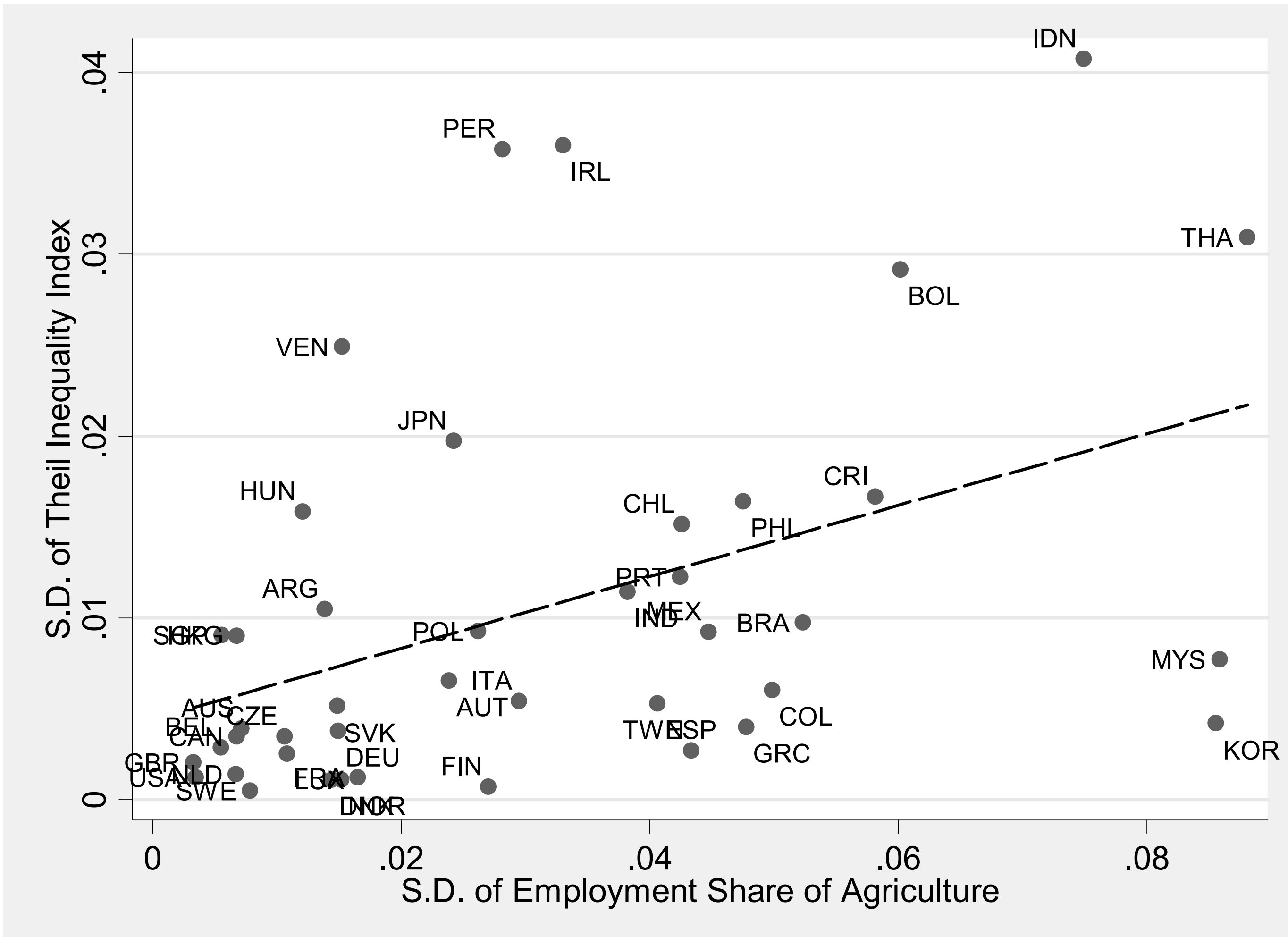


Figure 11 provides a global perspective on the change of inequality and its fluctuations. It presents the trend of annual average inequality since 1963 with their standard deviation across countries for each year, and numbers attached in the figure represent the number of countries included in the calculation of mean and standard deviation in each year. The change of average of inequality seems to accelerate in the 1980s and is especially high in the 1990s – not surprising considering the vast regime changes of that period. The same is true of cross-country variations: The mean of inequality level is stable up until 1980 and then increases with small fluctuations. Interestingly, the change of the variations across countries follows a similar pattern. It shrinks in the early 1980s and expands since then with considerable annual fluctuation in the 1990s. Table 3 presents the data on which the analysis presented so far is based.

Figure 11. Trend of Mean and Standard Deviation of Inequality across countries

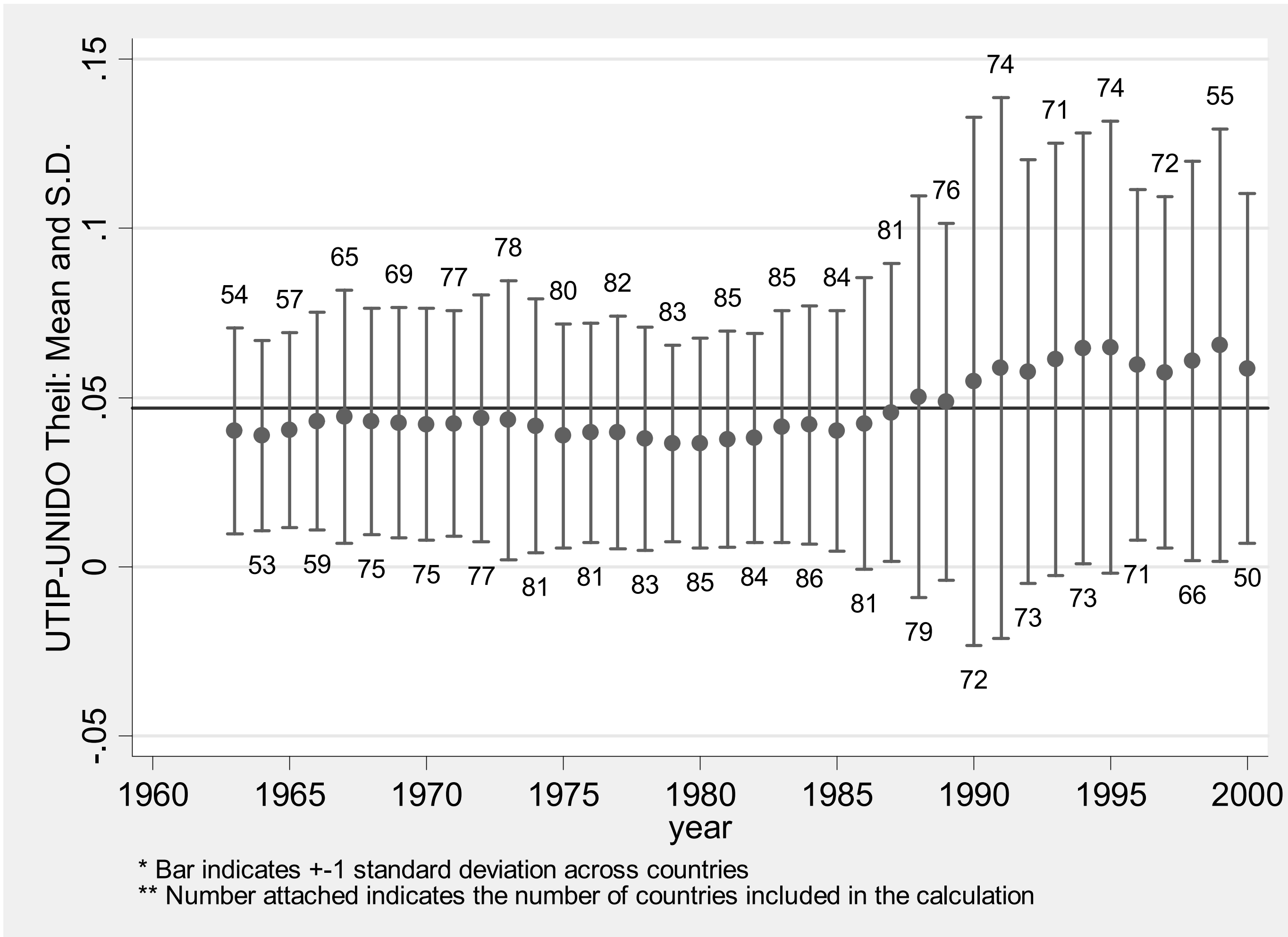


Table 3. Average and Standard Deviation of Employment Share by Sectors (Agriculture, Manufacturing, and Services) and UTIP-UNIDO Inequality for 1979-2000

Country code	AV_Agriculture	SD_Agriculture	AV_Manufacturing	SD_Manufacturing	AV_Services	SD_Services	AV_Theil	SD_Theil
SGP	0.8%	0.6%	25.4%	3.3%	64.0%	3.2%	0.04	0.009
HKG	1.0%	0.7%	25.5%	11.9%	64.6%	12.1%	0.014	0.009
USA	2.1%	0.3%	15.4%	2.4%	75.1%	3.2%	0.027	0.0012
GBR	2.2%	0.3%	18.1%	3.6%	71.2%	5.0%	0.017	0.002
LUX	2.9%	1.1%	17.9%	4.3%	66.0%	5.1%	0.015	0.0025
BEL	3.0%	0.6%	19.3%	2.8%	69.1%	3.8%	0.026	0.0029
SWE	3.5%	0.8%	19.3%	2.0%	70.2%	3.3%	0.004	0.0005
CAN	4.0%	0.7%	14.9%	1.9%	72.5%	2.9%	0.021	0.0034
DEU	4.0%	1.5%	25.4%	3.7%	59.8%	5.7%	0.013	0.0038
NLD	4.4%	0.7%	16.3%	2.4%	68.1%	3.9%	0.009	0.0014
DNK	5.2%	1.4%	18.5%	1.5%	69.0%	3.3%	0.007	0.0011
CZE	5.3%	1.1%	27.5%	0.6%	55.0%	2.0%	0.007	0.0035
AUS	5.5%	0.7%	14.8%	2.7%	69.7%	4.1%	0.013	0.0039
FRA	5.8%	1.7%	18.2%	2.7%	66.9%	5.2%	0.016	0.0012

NOR	6.0%	1.5%	15.3%	2.1%	69.9%	4.1%	0.01	0.0011
SVK	6.7%	1.5%	26.7%	1.3%	56.4%	3.6%	0.017	0.0051
HUN	7.5%	1.2%	24.3%	0.5%	58.7%	1.2%	0.02	0.0158
ITA	8.0%	2.9%	24.5%	2.3%	56.0%	4.6%	0.017	0.0054
JPN	8.9%	2.4%	20.4%	2.1%	59.8%	4.1%	0.04	0.0197
FIN	9.2%	2.7%	21.0%	1.8%	61.1%	5.0%	0.01	0.0007
ESP	10.9%	4.1%	19.6%	1.6%	53.0%	4.6%	0.027	0.0053
ARG	11.0%	1.4%	16.6%	3.0%	64.0%	5.8%	0.049	0.0105
IRL	12.1%	3.3%	20.2%	2.0%	57.3%	5.0%	0.044	0.036
TWN	13.2%	4.4%	30.9%	2.7%	45.9%	6.4%	0.015	0.0027
VEN	13.3%	1.5%	15.1%	2.0%	60.9%	3.3%	0.054	0.0249
PRT	15.2%	4.2%	22.5%	1.9%	51.8%	6.4%	0.041	0.0123
CHL	16.2%	4.3%	15.5%	2.0%	57.5%	5.5%	0.072	0.0151
AUT	16.4%	2.4%	19.1%	2.4%	56.2%	5.0%	0.019	0.0065
KOR	19.3%	8.7%	23.0%	2.8%	48.8%	8.9%	0.023	0.0042
GRC	22.7%	4.8%	17.6%	1.7%	50.1%	6.4%	0.032	0.004
MEX	22.8%	4.5%	19.1%	0.8%	49.9%	4.5%	0.03	0.0092
MYS	24.8%	8.6%	21.0%	5.0%	45.3%	3.1%	0.032	0.0077
POL	25.4%	2.6%	19.6%	1.3%	45.5%	3.5%	0.012	0.0092
CRI	26.2%	5.8%	17.5%	1.7%	47.8%	5.6%	0.04	0.0166
BRA	26.5%	5.2%	14.2%	1.2%	51.6%	6.8%	0.078	0.0097
COL	30.2%	5.0%	11.9%	0.9%	50.9%	4.7%	0.04	0.006
PER	34.6%	2.8%	11.5%	1.0%	48.5%	3.5%	0.089	0.0358
BOL	40.4%	6.0%	10.1%	2.0%	41.2%	4.4%	0.078	0.0291
PHL	44.1%	4.8%	10.0%	0.5%	40.5%	4.3%	0.066	0.0164
IDN	46.9%	7.5%	11.9%	2.1%	35.1%	3.9%	0.084	0.0407
THA	54.4%	8.8%	11.7%	2.5%	28.7%	5.1%	0.085	0.0309
IND	66.6%	3.8%	10.7%	1.3%	19.7%	1.8%	0.086	0.0114

* AV: Average, SD: Standard Deviation,
** Due to the small number of countries, 2001-2002 are omitted here.

5. The Behavior of Inequality: Regression Analyses

In this section, we examine the relationship between the UTIP-UNIDO inequality measures and several covariates to explain the behavior of inequality across countries and over time. Variables included in this analysis are GDP per capita in log term (ln_gdppc) and its square (ln_gdppc2), openness of the economy measured as the percentage of GDP (open), investment share of GDP (invest), the percentage of people over 65 (oldpop) and employment of manufacturing out of population (mnfemp).

The rationale for these covariates is straightforward. It has been a long tradition, going back to Kuznets, that inequality is related to income in quadratic form. In order to model this nonlinear behavior of inequality, including a quadratic term for GDP per capita is reasonable. However, Kuznets proposed the inverted-U curve but as Galbraith and Kum (2003) pointed out, a downward pattern with an upward tail at the high end is also possible. So the precise form is an object of empirical investigation. Openness and investment share in the GDP are employed to control for the economic factors that could affect manufacturing industry. A negative influence of investment share on inequality is expected but the effect of openness is not clear *a priori*. The employment share of manufacturing in population and the share of the elderly in total population reflect the demographic structure in a society. The manufacturing employment share is expected to have negative sign, and for the elderly population share a positive sign is expected.

The specification of the model is:

(1) $$Y_{ij} = \beta_1 + \beta_2 X_{2ij} + \beta_3 X_{3ij} + \beta_4 X_{4ij} + \beta_5 X_{5ij} + \beta_6 X_{6ij} + \beta_7 X_{7ij} + \varsigma_{1i} + \varsigma_{2j} + \varepsilon_{ij}$$

where X's are the covariates; X_2 is GDP per capita in log term and X_3 is its square term. X_4 is openness, X_5 investment share, X_6 manufacturing employment share, X_7 is old population share. Y is an inequality index, UTIP-UNIDO Theil index. The peculiar point in this specification is that there are two additional error terms, ς_{1i} and ς_{2j}, for country-specific (i) effects and year-specific (j) effects in the model. Since it is unrealistic that the behavior of inequality in each country is deterministic, statistical models which allow the effects of either country or year to vary by specifying the two-way error-components have been heavily specified in the literature. Depending on the assumption about the features of these error terms, fixed effects models and random effects models could be estimated. The residual error ε_{ij} includes both the interaction between year and country and any other effect specific to country i in year j.

We assume that ς_{1i} and ς_{2j} are uncorrelated with the residual error (ε_{ij}) and they are all normally distributed. Random effects models also assume that the random effects have zero means and are correlated with neither each other nor the covariates in the model, which leads to the efficient estimator. Fixed effects models, however, do not take the orthogonality of ς_{1i} and ς_{2j} are uncorrelated with the included covariates. For this, it is necessary to treat them as additional constant term for each country or/and year for consistent estimator.[7] We estimate both fixed effects and random effects models with the Hausman specification test to compare these estimators.[8] World Development Indicators (2007) and Penn World Table 6.2 are the data sources for these variables and the dependent variable is UTIP-UNIDO inequality measure. We use 2,607 observations for 86 countries that have at least 20 annual observations in the 1963-2002 period.

---

7 This is why the fixed effects model is often termed as the least-squares dummy variable (LSDV) model.

8 For more information on the fixed effects model and random effects model, see Wooldridge (2002)

Table 4. Results of Regression Estimates

	(1) OLS Ln_Theil	(2) FIX-ONE Ln_Theil	(3) RANONE Ln_Theil	(4) FIXTWO Ln_Theil	(5) RANTWO Ln_Theil	(6) RANCOEF Ln_Theil
ln_gdppc2	-0.036	-0.174	-0.152	-0.144	-0.133	-0.163
	(2.63)**	(7.62)**	(6.95)**	(6.48)**	(6.38)**	(3.39)**
ln_gdppc	0.758	3.09	2.719	2.312	2.164	2.79
	(3.21)**	(7.79)**	(7.18)**	(5.87)**	(5.94)**	(3.35)**
open	0.001	0.001	0.001	0.0001	0.0001	0.002
	(5.33)**	(2.74)**	(3.32)**	(0.38)	(0.49)	(4.64)**
Invest	-0.014	-0.008	-0.011	-0.0004	-0.003	-0.009
	(7.85)**	(3.91)**	(5.19)**	(0.2)	(1.57)	(4.43)**
mnfemp	-0.252	-0.268	-0.314	-0.159	-0.208	-0.253
	(9.32)**	(8.10)**	(9.99)**	(4.93)**	(6.81)**	(6.78)**
oldpop	-0.076	0.064	0.035	0.028	0.006	0.120
	(14.31)**	(5.65)**	(3.41)**	(2.43)*	(0.6)	(7.88)**
Constant	-5.651	-16.449	-14.491	-12.22	-11.328	-14.964
	(5.69)**	(9.59)**	(8.89)**	(7.03)**	(7.14)**	(4.07)**
Obs.	2607	2607	2607	2607	2607	2607
countries		86	86	86	86	86
R-squared	0.38	0.07		0.17		
log likelihood			-1658.37		-1571.43	-1479.42
RE_SD(country)			0.717		0.636	7.727
RE_SD(year)					0.191	
RE_SD(resid)			0.425		0.404	0.375
RE_SD(gdppc)						0.936

Absolute value of t or z statistics in parentheses.
* significant at 5%; ** significant at 1%

Table 4 reports the results of the estimation. Column 1 is the result of a pooled-OLS model for reference and column 2-5 are the results of fixed effects model and random effects model with country and year effects respectively. In the every specification model, the coefficients for both GDP per capita and its square term are significant at 5% level indicating an inverted-U shape. With regard to the range of GDP per capita, however, as the model under consideration is moving from one-way fixed effects model (column 2) to two-way random effects model (column 5) in the Table 4, the turning points of inverted-U curve gradually shifts to the left, which means that more and more predicted values are placed on the downward portion of the curve.

The sign of the estimated coefficient for manufacturing employment share is negative as expected in all models with variations of assumptions about country-specific and year-specific effects (fixed or random). In fact, the effect of manufacturing employment share on inequality is robust in terms of the magnitude and statistical significance across models. The estimated coefficient of the investment share is negative in all models but not significant when year-effects are included. Elderly population share is estimated to have a positive effect on inequality, but the magnitudes of the estimated coefficients across models are decreasing. Openness is estimated to have positive and statistically significant effects on inequality only when country-specific effects are considered, but the magnitude is very small.

The random effects models provide additional parameter estimates, that is, estimated residual standard deviation (RE_SD). In a two-way random effects model (column 5), the estimated residual standard deviation between countries (RE_SD(country)) is 0.64, which is smaller than that of one-way random effects model (0.72). Also the estimated residual standard deviation between years (RE_SD(year)) is 0.19, and the remaining residual variability, not due to additive effects of countries and years, is 0.4 (RE_SD(resid)). Based on these estimated residual standard deviations, the intra-class correlation for countries given a year is calculated as 0.67 and the intra-class correlation for years given a country is as 0.06. That means there is a much higher correlation over years within a country than that over countries within a year, given the covariates. In other words, the change of inequality within country is relatively small while the differences of inequality between countries are relatively large. This point matches the conventional wisdom that the variations of inequality across countries are larger than those are from within country through years (Li-Zou-Squire, 1997).

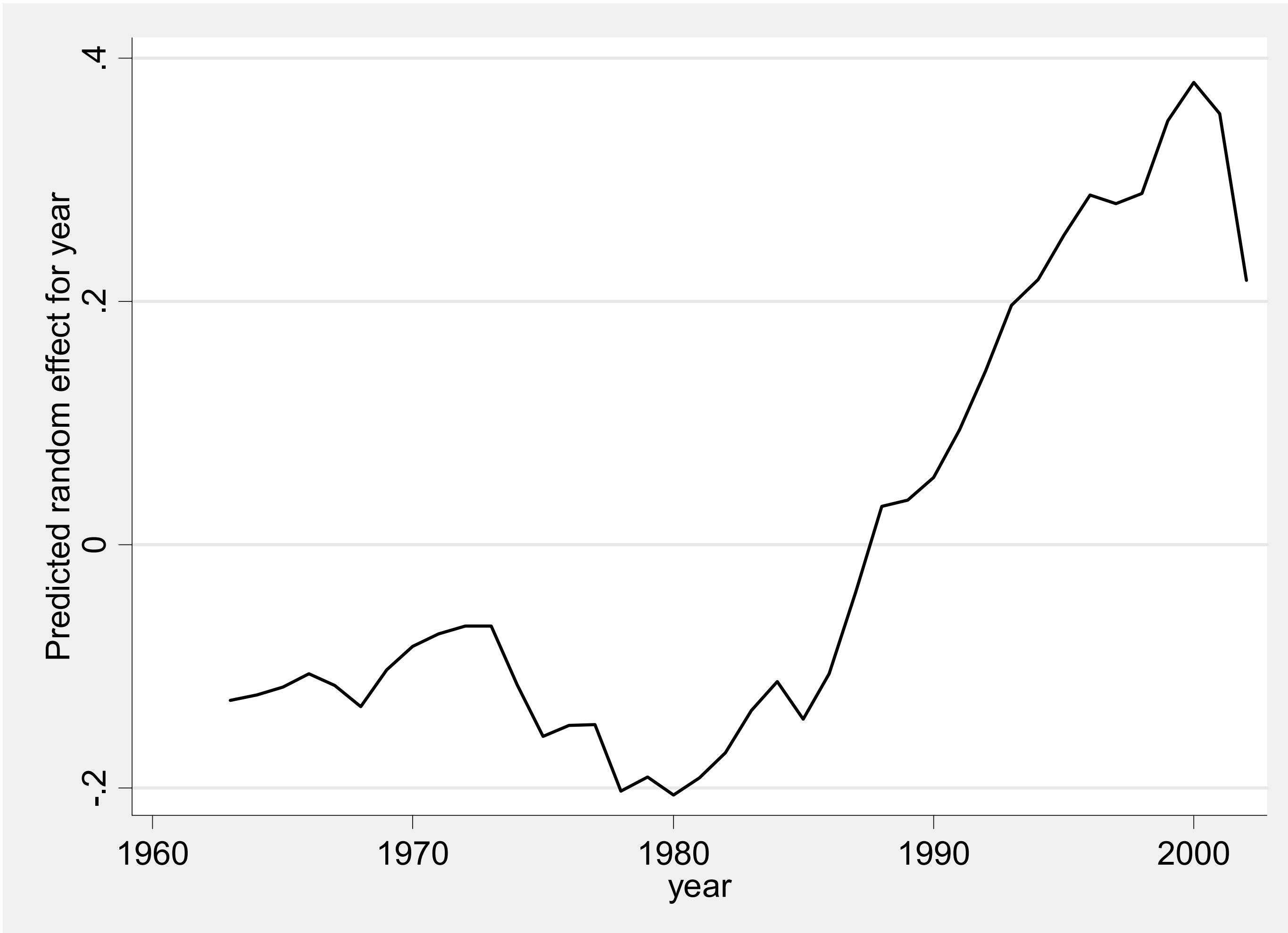
The estimated random effects for countries and years are listed in Table 5. We can see that the prediction of random effects for the Sweden (SWE) is -1.677, which means that after controlling for covariates specified in the model, Sweden (SWE) is a country with lowest pay inequality relative to the average across countries. On the other hand, the prediction of random effects for 2000 is 0.380, which means, after controlling for covariates specified in the model, 2000 is a year with highest pay inequality than the average over the nearly 40 year time span.

Table 5 provides another interesting point of time effects on inequality. As can be seen in the Figure 13, time effects of inequality are independent from other covariates in the model as well as country-specific effects. Thus, their movement indicates other macro factors that influence the inequality in a global sense. Table 5 shows that the time effects decreased up until 1980 but after that it turned up and kept increasing. This had already been found by Galbraith and Kum (2003). But after 2000, this trend is reversed. Evidently the global financial climate improved dramatically after 2000 for poorer countries and for low-paid working people within them.

Table 5. Estimated Random Effects for countries and year

Country code	random effects for country	Country code	random effects for country	Year	random effects for year
SWE	-1.677	ERI	0.080	1963	-0.128
DZA	-1.319	BEL	0.096	1964	-0.124
POL	-1.238	ESP	0.134	1965	-0.117
DNK	-1.215	SYR	0.141	1966	-0.106
CZE	-1.127	CYP	0.158	1967	-0.116
HUN	-0.910	USA	0.169	1968	-0.133
NLD	-0.889	LKA	0.169	1969	-0.103
NOR	-0.863	HND	0.175	1970	-0.083
MLT	-0.819	CIV	0.192	1971	-0.074
MAC	-0.789	RUS	0.237	1972	-0.067
AUS	-0.772	ZWE	0.240	1973	-0.067
FIN	-0.758	BGR	0.275	1974	-0.115
HKG	-0.675	VEN	0.286	1975	-0.158
DEU	-0.643	PHL	0.286	1976	-0.148
NIC	-0.639	URY	0.294	1977	-0.148
MEX	-0.558	ISL	0.306	1978	-0.203
IRQ	-0.533	PRT	0.331	1979	-0.191
SEN	-0.502	PAN	0.339	1980	-0.206
BGD	-0.490	CHL	0.340	1981	-0.192
IRN	-0.489	JPN	0.354	1982	-0.171
EGY	-0.411	GTM	0.441	1983	-0.137
GBR	-0.390	ISR	0.473	1984	-0.112
LUX	-0.375	TUN	0.481	1985	-0.144
FRA	-0.372	IND	0.483	1986	-0.106
NZL	-0.342	MAR	0.495	1987	-0.039
ITA	-0.317	IDN	0.495	1988	0.032
AUT	-0.263	ZAF	0.496	1989	0.037
NGA	-0.239	KEN	0.520	1990	0.055
CAN	-0.207	PNG	0.526	1991	0.094
MDG	-0.199	JOR	0.529	1992	0.143
KOR	-0.159	BRB	0.556	1993	0.197
COL	-0.146	CMR	0.582	1994	0.218
TUR	-0.130	YUG	0.594	1995	0.255
CRI	-0.129	TZA	0.635	1996	0.287
HTI	-0.126	MWI	0.654	1997	0.280
FJI	-0.120	MUS	0.705	1998	0.289
MYS	-0.084	DOM	0.706	1999	0.349
ECU	-0.079	SWZ	0.737	2000	0.380
PAK	-0.063	SGP	0.788	2001	0.355
IRL	-0.040	GHA	0.900	2002	0.218
GRC	-0.027	JAM	1.030		
SLV	-0.023	TTO	1.379		
BOL	0.018	KWT	2.323		

Figure 13. Predicted random effects, by year



So far we have investigated the behavior of pay inequality by using two-way random effects (intercepts) model, which assume that the country-specific regression lines are parallel with common time-specific effects. Thus we have allowed the country-specific effects and year-specific effects to vary randomly but those effects are still part of error terms and could not affect the major relationship between inequality and GDP per capita in the model. As a final specification, we incorporate the idea that countries would differ in their overall rate of inequality change according to their GDP per capita. In this case, the relationship between inequality and GDP per capita would vary by country and we can specify the model as follows:,

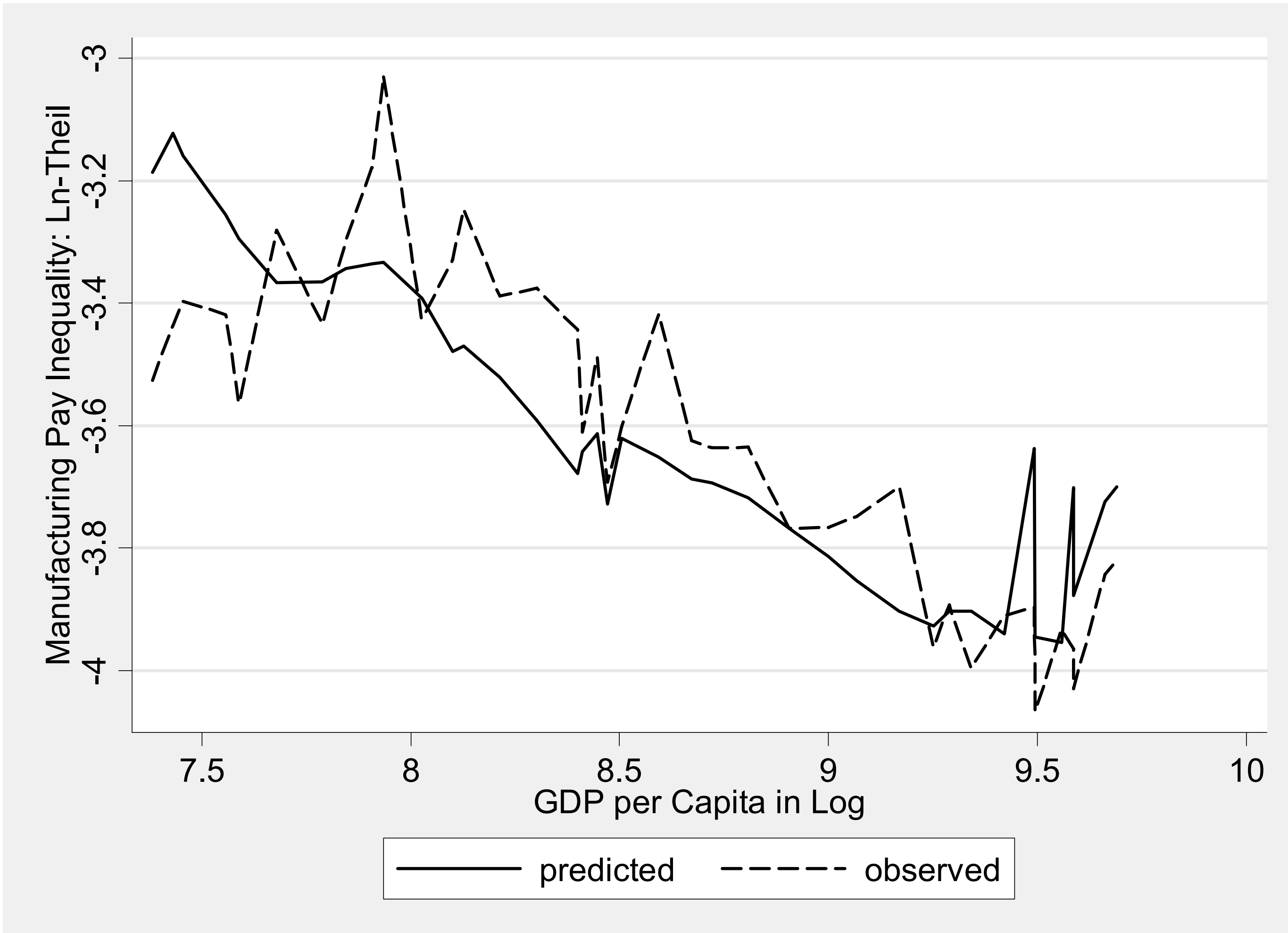
$$(2)\quad Y_{ij} = (\beta_1 + \varsigma_{1i}) + (\beta_2 + \varsigma_{2i})X_{2ij} + \beta_3 X_{3ij} + \beta_4 X_{4ij} + \beta_5 X_{5ij} + \beta_6 X_{6ij} + \beta_7 X_{7ij} + \varepsilon_{ij}$$

where a random slope coefficient for GDP per capita (ς_{2j}) is included. With this specification, we can assume a country-specific slope of GDP per capita.

The estimated coefficients are reported in the last column (column 6) of Table 4. The sign and the magnitude of the estimated coefficients are similar to those of the previous models and all are statistically significant, which indicate the robustness of the relationship between inequality and the covariates in the models. Compared with the one-way random effects model (column 3) and two-way random effects model (column 5), the random intercept and slope model (column 6) has improved log likelihood from -1658 and -1571 to -1479. This indicates that the inclusion of random slopes makes the model fit better. Also the estimated random-slope standard deviation (RE_SD(gdppc)) is 0.936, and the estimated residual standard deviation (RE_SD(resid)) has decreased from 0.425 (column 3) and 0.404 (column 5) to 0.375 reflecting the better fit again of the inequality change trajectories with random slope.

Figure 14 plots the predicted values from the random intercept and random coefficient model against observed inequality for Korea. Although the prediction does not cover the annual fluctuations, it seems to capture the overall trend of inequality quite well. For instance, the Pearson correlation coefficient between the observed and predicted inequality from random intercept and random slope model is 0.91.

Figure 14. Predicted and Observed Inequality against GDP per Capita (in log)



6. Conclusion

This paper has presented and summarized new measures of inequality within-countries for the world economy in the years 1963-2002, and has shown the relationship between those measures and indicators of structural change for 42 countries during this period. As a general conclusion, measures of inequality in manufacturing pay appear closely related to the share of agricultural employment, suggesting that the former are a sensible indicator of the overall state of economic development. Further, it is clear that countries with high inequality experience much more variability of inequality than countries with low inequality – an indication of the institutional strength of the latter. Finally, previous findings of a Kuznets relation and a global pattern in the evolution of inequality are confirmed, though after 2000 the long period of rising global inequality appears to have come to an end.

References

Berman. E. 2000. "Does Factor-Biased Technological Change Stifle International Convergence? Evidence from Manufacturing." National Bureau of Economic Research. Working Paper NO. 7964. October.

Conceição P., Galbraith J., Bradford P. 2001. "The Theil Index in Sequences of Nested and Hierarchical Grouping Structures: Implications for the Measurement of Inequality Through Time, With Data Aggregated at Different Levels of Industrial Classification," Eastern Economic Journal, **27**, 491-514

Galbraith. J. K.and H. Kum. "Inequality and Economic Growth: A Global View based on Measures of Pay." CESifo Economic Studies. Vol. 4. 2003.

Kuznets. S. 1955. "Economic Growth and Income Inequality." American Economic Review. Vol. 45. 1–28.

Li. H., H. Zou and L. Squire. 1998. "Explaining International and Intertemporal Variations in Income Inequality." Economic Journal. Vol. 108. 26–43.

Pyatt. G. 1976. "On the interpretation and disaggregation of Gini coefficients." Economic Journal. Vol. 86. 243-255.

Rodrik. D. 1999. "Democracies pay higher wages." Quarterly Journal of Economics, Vol. 114. 707-738.

Theil. H. 1972. Statistical Decomposition Analysis: with Application to the Social and Administrative Science. Amsterdam: North Holland.

Wooldridge. J. M. 2003. Introductory Econometrics: Modern Approach. Second ed. Thompson Learning-South-Western.

UNIDO (United Nations International Development Organization) Industrial Statistics Database (2005).

Groningen Growth and Development Center. 2007. 10-Sector Database and 60-Industry Database. University of Groningen. http://www.ggdc.net/

Center for International Comparisons, Penn World Table 6.2. University of Pennsylvania. http://pwt.econ.upenn.edu/

World Bank. World Development Indicator 2007

University of Texas Inequality Project (UTIP). http://utip.gov.utexas.edu

Appendix

- **UTIP-UNIDO Theil Index (countries with N>20)**

Country	Mean Theil	Std. Dev.	N	Country	Mean Theil	Std. Dev.	N
AUS	0.011	0.004	39	JOR	0.082	0.023	37
AUT	0.018	0.005	37	JPN	0.035	0.017	39
BEL	0.026	0.002	35	KEN	0.082	0.021	40
BGD	0.031	0.021	28	KOR	0.028	0.007	39
BGR	0.022	0.027	40	KWT	0.240	0.112	35
BOL	0.066	0.032	31	LKA	0.059	0.019	22
BRB	0.058	0.016	28	LUX	0.014	0.003	38
CAN	0.019	0.004	39	MAC	0.010	0.004	25
CHL	0.060	0.022	38	MAR	0.083	0.020	26
CIV	0.064	0.014	22	MDG	0.043	0.023	22
CMR	0.138	0.091	25	MEX	0.028	0.009	31
COL	0.038	0.005	38	MLT	0.016	0.007	39
CRI	0.038	0.016	22	MUS	0.071	0.026	32
CYP	0.041	0.013	40	MWI	0.092	0.049	32
CZE	0.007	0.003	33	MYS	0.034	0.008	33
DEU	0.012	0.003	38	NGA	0.043	0.018	28
DFA	0.003	0.001	23	NIC	0.023	0.008	21
DNK	0.006	0.001	36	NLD	0.010	0.002	38
DOM	0.079	0.026	23	NOR	0.009	0.001	39
DZA	0.015	0.018	28	NZL	0.018	0.010	36
ECU	0.050	0.028	37	PAK	0.046	0.016	30
EGY	0.035	0.025	36	PAN	0.065	0.019	37
ERI	0.070	0.023	35	PHL	0.062	0.015	35
ESP	0.034	0.010	38	PNG	0.087	0.026	27
FIN	0.011	0.001	38	POL	0.011	0.008	31
FJI	0.044	0.031	27	PRT	0.042	0.011	32
FRA	0.016	0.001	21	RUS	0.022	0.028	38
GBR	0.016	0.002	34	SEN	0.047	0.029	29
GHA	0.113	0.041	28	SGP	0.056	0.026	40
GRC	0.032	0.004	36	SLV	0.052	0.029	29
GTM	0.098	0.076	26	SWE	0.004	0.001	38

HKG	0.013	0.008	30	SWZ	0.108	0.048	26
HND	0.059	0.030	26	SYR	0.076	0.062	36
HTI	0.053	0.020	21	TTO	0.155	0.082	26
HUN	0.016	0.013	38	TUN	0.084	0.054	28
IDN	0.087	0.035	33	TUR	0.043	0.021	38
IND	0.078	0.019	39	TWN	0.016	0.003	25
IRL	0.034	0.030	38	TZA	0.078	0.027	29
IRN	0.039	0.027	38	URY	0.050	0.015	26
IRQ	0.031	0.012	27	USA	0.025	0.003	38
ISL	0.037	0.027	29	VEN	0.053	0.023	34
ISR	0.046	0.021	39	YUG	0.034	0.037	35
ITA	0.020	0.007	34	ZAF	0.059	0.007	36
JAM	0.122	0.060	27	ZWE	0.057	0.033	34