INCOME DISPARITIES, ECONOMIC GROWTH,
AND DEVELOPMENT AS A THRESHOLD

SHERIF KHALIFA AND SHERINE EL HAG*

California State University, Fullerton

Galor and Moav (2004) argue that in the early stages of development, physical capitalaccumulation is the primary source of economic growth. Thus, inequality enhances growth by channeling resources towards individuals whose marginal propensity to save is higher. In later stages of development, physical capital is replaced by human capital as the engine of growth. Accordingly, equality alleviates the adverse effects of credit constraints on human capital accumulation and prompts the growth process. This paper attempts to test empirically the finding that the impact of income inequality on economic growth depends on the development stage. A threshold estimation technique, developed by Hansen (1999), is utilized for a panel of 70 countries for the period between 1970 and 1999. The estimation suggests that there is a statistically significant threshold income per capita, below which the coefficient on the relationship between inequality and growth is significantly negative and above which the estimate is positive, but not statistically significant.

Keywords: Income Inequality, Economic Growth
JEL classification: D9, D31

1. INTRODUCTION

Income distribution and its effect on economic outcomes have always been a source of concern for economists. In this context, there are two streams of literature. One argues that income inequality is propitious to economic performance, while the other concludes that the prevalent disparities call for an intervention to achieve the desired outcomes. According to Kuznets (1955), these attempts struggled in a “field of study that has been plagued by looseness in definitions, unusual scarcity of data, and pressure of strongly held opinions”.

* We would like to thank an anonymous referee for very useful comments. All remaining errors are our own.
On one hand, a vast literature argues that greater egalitarian conditions are a prerequisite for economic growth, and that inequality adversely affects the overall performance of the economy. For instance, Barro (2000) shows that redistribution from the rich whose marginal productivity is low to the poor whose marginal productivity is high, but cannot invest in human capital more than their endowment due to capital market imperfections, would enhance productivity and growth. Credit market imperfections cause the investment opportunities to depend on the individuals’ assets and incomes. Thus, a redistribution of assets and incomes from rich to poor tends to increase investments, and accordingly a reduction in inequality enhances economic growth.

From a political economy point of view, Persson and Tabellini (1994) show that in more inegalitarian economies, a majority of voters prefer higher level of redistribution which reduces the incentives for investment, adversely affecting growth. Alesina and Rodrick (1994) demonstrate that the more equitable the economy, the better endowed the median voter with capital, the lower the level of capital taxation and the higher is the economy’s growth. Finally, Perotti (1992) emphasizes that social disparities motivate disruptive and destabilizing activities, and argues that redistributive policies reduce social tension.

Another stream of literature disputes the previous findings and asserts that the skewness of income distribution is conducive to economic performance. In this context, studies of consumption and saving behavior proposed a channel in which inequality has a stimulating effect on growth. For instance, Carroll (2000) finds that the marginal propensity to save of the rich is higher than that of the poor. The implication is that if the growth rate of income is proportional to aggregate saving, more inegalitarian economies, where wealth is concentrated in the hands of the rich, are bound to grow faster.

As an attempt to reconcile these two streams, Galor and Moav (2004) provide a unified approach arguing that in the early stages of development, physical capital accumulation is the primary source of economic growth. Hence, inequality enhances growth by channeling resources towards individuals whose marginal propensity to save is higher. In later stages of development, physical capital is replaced by human capital as the engine of growth. Accordingly, equality alleviates the adverse effects of credit constraints on human capital accumulation and prompts the growth process. Therefore, inequality enhances growth in developing countries, and hinders growth in developed ones.

This paper attempts to test empirically the finding that the impact of income inequality on economic growth depends on the development stage, as proposed by Galor and Moav (2004). A threshold estimation technique for a non-dynamic panel, developed by Hansen (1999), is utilized for 70 countries for the period between 1970 and 1999. The threshold regression model allows the level of Gross Domestic Product GDP per capita to determine the existence and significance of a threshold level in the relationship between income inequality and economic growth. The estimation suggests that there is a statistically significant threshold income per capita, below which the coefficient on the
relationship between income inequality and economic growth is significantly negative and above which the estimate is positive, but not significant.

The remainder of the paper is organized as follows: section 2 includes a detailed survey of the empirical literature, section 3 includes the estimation, section 4 concludes, section 5 is the data appendix. References, tables, and figures are included thereafter.

2. LITERATURE

The contribution of this paper is highlighted in comparison to the previous literature on the empirical estimation of the relationship between income inequality and economic growth. This section includes a detailed comprehensive survey of the literature in terms of the data used, the econometric techniques adopted, and the conclusions in these studies.

Due to the lack of panel data on income inequality, earlier studies in this literature used cross sectional ordinary least squares analyses of cross country data. For instance, Alesina and Rodrick (1994) study the relationship between politics and economic growth in a model with distributive conflict. The analysis for the 1960-1985 period, indicate that income and land inequality are negatively correlated with subsequent growth. The results also reject the hypothesis that this relationship is different across democracies and non democracies. Persson and Tabellini (1994) use historical data of developed countries and show that the coefficient of income inequality, measured as the share of the income of the top 20%, on growth is significantly negative. Including time dummies cause this coefficient to be insignificant. Examining post war data, from a broader cross section of countries, show that the increase in the income of the middle class enhances growth only in democratic countries. Perotti (1996) uses a sample of 67 countries around 1960, to show that the income share of the middle class, as a measure of equality, is positively associated with growth. Intercontinental variation in income distribution, however, accounts for a substantial part of the variation in this result. The study also shows that the coefficient of equality is much higher and significant in democracies and in rich countries, but insignificant in non democracies and in poor ones. In addition, an increase in the share of the middle class decreases sociopolitical instability, reduces fertility, and induces more investment in education, which all lead eventually to higher growth. Alesina and Perotti (1996) use a cross section of 71 countries, for the period 1960-1985, to show that income inequality hinders economic growth. In a bivariate simultaneous equation model, they find that a rich middle class reduces sociopolitical instability and that instability depresses investment and growth. They concluded that a healthy middle class is conducive to capital accumulation because it creates conditions of social stability. The introduction of regional dummies reduces the coefficient of the middle class, but it remains significant.

The inequality data used in these studies are considered deficient in their quality, their comparability over time and across countries, and their geographical and temporal
coverage. Deininger and Squire (1996) compiled a dataset on inequality that is more consistent and comprehensive. They filtered out the observations that satisfied minimum standards of quality such that: the data must be based on household surveys, be representative of the population at the national level, and have coverage of all sources of income. Since this data has a time series dimension for enough countries, more advanced panel estimation techniques became possible to adopt in all studies thereafter.

In this context, Deininger and Squire (1998) use their own compiled data for the period 1960-1992 and show that initial inequality affects future growth negatively. However, the coefficient on inequality ceases to be significant once regional dummies are introduced. Their results also reveal that initial inequality of land distribution tends to reduce long term growth, and the significance of the results is robust to the addition of regional dummies. The coefficient of land inequality on future growth is found significant only in a sample of developing countries, while the variable is insignificant if only high income countries are considered. Benhabib and Speigel (1998) investigate whether ancillary variables, such as income distribution, affect economic growth. For the period 1960-1980, the GMM estimation shows that the effect of income inequality is insignificant. They conclude that “none of the non financial ancillary variables enter significantly as determinants of growth after accounting for disparities in rates of factor accumulation”. Barro (2000) applies a three stage least squares, over the period 1960-1990, and finds that income inequality has no significant relation with subsequent economic growth. When the effect of the Gini coefficient on economic growth is allowed to depend on the real GDP per capita, the estimation implies that the effect of inequality on growth is negative for values of GDP per capita below $2070 and then becomes positive. When the inequality measure is based on quintile-shares data, the effect of inequality on growth is negative when GDP per capita is less than $1473, and positive otherwise.

It is obvious that studies that used the Deininger and Squire (1996) data found an insignificant coefficient of inequality on growth; especially, when additional explanatory variables or regional dummies are included. This indicates an omitted variable bias. Unlike earlier studies that consistently found a negative relation between inequality and growth, the latter papers found a significantly negative coefficient only in subsamples of poor and developing countries. This induced other studies to adopt advanced econometric techniques to resolve the contradictions in the literature.

For instance, Forbes (2000) applies the Arellano and Bond (1991) technique, on data covering the period 1966-1995, to control for any time-invariant omitted variables. An increase in a country’s level of income inequality is found to have a significant positive relationship with subsequent growth. This relationship is highly robust across samples, variable definitions and model specifications, with the one caveat that it may not apply to very poor countries. Banerjee and Duflo (2003) examine the data without imposing a linear structure that previous studies adopted. Using nonparametric methods, they find that changes in inequality, in any direction, are associated with reduced growth in
inequality. They also find a negative relationship between growth rates and inequality lagged one period.

The empirical contribution of this paper to the existing literature is obvious in terms of the data set used, the coverage period, the econometric technique, and the conclusions of the paper. Contrary to the previous literature, this paper relies on a new data set that was not used before. This data is compiled by the University of Texas Inequality Project, while most of the recent literature relied on the Deininger and Squire (1996) data. Scholars expressed their unease about the quality of this dataset as the coverage is sparse and unbalanced. This implies that studies attempting to assess the time trend of inequality must either be affected by the bias that may be associated, restrict their attention to a subset of these data, or attempt to fill in the gaps by interpolation. The University of Texas Inequality Project compiled an alternative global inequality dataset, which has more observations than in the Deininger and Squire dataset, and are based on more accurate sources, both through time and across countries.¹

Contrary to the previous literature, this paper uses a new econometric technique which is the threshold estimation developed by Hansen (1999). The advantage of this technique is that it allows for the endogenous determination of the estimate and significance of a threshold development level, besides the coefficients below and above the threshold. Unlike previous studies that estimated different coefficients for poor and rich countries, this is achieved without splitting the sample. Previous studies had to impose an arbitrary classification scheme to distinguish between developing and developed countries, and split the sample in order to run separate regressions for each subsample to find different coefficients for every group of countries. This is especially critical since the paper is attempting to empirically estimate the finding in Galor and Moav (2004) that the relationship between income inequality and economic growth depends on the development stage.

The use of the new technique leads to new conclusions as well. It is obvious that the previous literature, that used panel estimation, found an insignificant coefficient of inequality on growth. The coefficient becomes significantly negative in a subsample of poor countries only. In this paper, the coefficient is found significantly negative for countries whose income per capita is below the estimated threshold, without splitting the sample.

3. ESTIMATION

In this section, the finding that the relationship between income inequality and economic growth depends on the development stage, is tested empirically using the threshold estimation technique developed in Hansen (1999). The econometric model is

¹ Detailed data description is included in the appendix.
typical to that used in the literature to estimate the effect of income inequality on economic growth. The specification, as in Perotti (1996) and Forbes (2000), estimates the growth rate as a function of lagged income inequality, a lagged measure of human capital, and lagged market distortions. The model is typical to that used in previous studies so that any discrepancy between this paper and previous work cannot be explained by model specification. The threshold estimation model is, thus, given by

\[
\begin{align*}
\text{Growth}_it = \begin{cases} 
\mu_i + \beta_1 \text{Gini}_it-1 + \phi_1 \text{Education}_it-1 + \phi_2 \text{Distortions}_it-1 + \epsilon_it, & \text{if } \text{GDP}_it-1 \leq \sigma \\
\mu_i + \beta_2 \text{Gini}_it-1 + \phi_1 \text{Education}_it-1 + \phi_2 \text{Distortions}_it-1 + \epsilon_it, & \text{if } \text{GDP}_it-1 > \sigma
\end{cases}
\end{align*}
\]

(1)

where the subscript \(i\) indexes the country, and the subscript \(t\) indexes time. The dependent variable \(\text{Growth}_it\) denotes the growth rate of GDP per capita in country \(i\) in year \(t\). The threshold variable \(\text{Gini}_it-1\) is a measure of the Gini coefficient in country \(i\) in year \(t-1\). The variable \(\text{Education}_it-1\) is a measure of educational attainment in country \(i\) in year \(t-1\). The variable \(\text{Distortions}_it-1\) is a measure of market distortions in country \(i\) in year \(t-1\), and is the threshold variable determining the stage of development. Usually, initial real GDP per capita is included as an independent variable in the previous literature to test for convergence. If included, the equation contains a lagged endogenous variable which is the income term. As the Hansen’s (1999) technique is developed for a non-dynamic panel, lagged real GDP per capita is excluded from this regression.

In this context, the threshold GDP per capita determines whether the coefficient on the Gini coefficient is positive or negative. In this context, the observations are divided into two regimes depending on whether the threshold variable \(\text{GDP}_it-1\) is smaller or larger than the threshold \(\sigma\). The regimes are distinguished by differing regression slopes, \(\beta_1\) and \(\beta_2\). Following Hansen (1999), another way of writing the equation of interest is

\[
\begin{align*}
\text{Growth}_it = \mu_i + \beta_1 \text{Gini}_it-1 I(\text{GDP}_it-1 \leq \sigma) + \beta_2 \text{Gini}_it-1 I(\text{GDP}_it-1 > \sigma) \\
+ \phi_1 \text{Education}_it-1 + \phi_2 \text{Distortions}_it-1 + \epsilon_it,
\end{align*}
\]

(2)

where \(I(.)\) is the indicator function. A balanced panel annual data is used for 70 countries and cover the period from 1970 to 1999. A Gini coefficient compiled by the University of Texas Inequality Project is used as a proxy for income inequality. The average years of total education in the population aged over 15, from Barro and Lee data on educational attainment, is used as a measure of human capital. Finally, real GDP per capita, and market distortions are extracted from the Penn World Tables 6.2. Detailed data description is included in the appendix. Summary statistics of the variables used in the estimation are provided in table 1.
Table 1. Summary Statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Minimum</th>
<th>25% Quantile</th>
<th>Median</th>
<th>75% Quantile</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Growth$_t$</td>
<td>-1.004230</td>
<td>-0.002482</td>
<td>0.021078</td>
<td>0.043438</td>
<td>0.463332</td>
</tr>
<tr>
<td>Education$_t$</td>
<td>0.063400</td>
<td>0.750200</td>
<td>1.474800</td>
<td>2.396000</td>
<td>5.742000</td>
</tr>
<tr>
<td>Female Education$_t$</td>
<td>0.201600</td>
<td>1.112000</td>
<td>2.248000</td>
<td>1.221000</td>
<td>1.494600</td>
</tr>
<tr>
<td>Male Education$_t$</td>
<td>0.580800</td>
<td>1.251000</td>
<td>1.712400</td>
<td>1.462000</td>
<td>2.028600</td>
</tr>
<tr>
<td>Gini$_t$</td>
<td>24.069090</td>
<td>36.424790</td>
<td>42.397830</td>
<td>47.224380</td>
<td>58.975360</td>
</tr>
<tr>
<td>GDP$_t$</td>
<td>474.417775</td>
<td>3359.168884</td>
<td>6322.40258</td>
<td>15183.425690</td>
<td>64336.281600</td>
</tr>
<tr>
<td>Distortions$_t$</td>
<td>15.361997</td>
<td>59.927724</td>
<td>77.666764</td>
<td>100.161582</td>
<td>2654.164895</td>
</tr>
</tbody>
</table>

Table 2. Tests for Threshold Effects

<table>
<thead>
<tr>
<th>Test</th>
<th>Regression 1</th>
<th>Regression 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test for Single Threshold</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$F_1$</td>
<td>87.026828</td>
<td>86.901446</td>
</tr>
<tr>
<td>$P$-value</td>
<td>0.000000</td>
<td>0.000000</td>
</tr>
<tr>
<td>10% critical value</td>
<td>22.349967</td>
<td>20.341321</td>
</tr>
<tr>
<td>5% critical value</td>
<td>29.048133</td>
<td>27.382181</td>
</tr>
<tr>
<td>1% critical value</td>
<td>44.483860</td>
<td>42.191501</td>
</tr>
<tr>
<td>Test for Double Threshold</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$F_2$</td>
<td>51.071527</td>
<td>53.855204</td>
</tr>
<tr>
<td>$P$-value</td>
<td>0.003333</td>
<td>0.000000</td>
</tr>
<tr>
<td>10% critical value</td>
<td>20.043570</td>
<td>21.017158</td>
</tr>
<tr>
<td>5% critical value</td>
<td>25.606569</td>
<td>25.301465</td>
</tr>
<tr>
<td>1% critical value</td>
<td>42.554054</td>
<td>33.809370</td>
</tr>
<tr>
<td>Test for Triple Threshold</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$F_3$</td>
<td>11.739752</td>
<td>10.821065</td>
</tr>
<tr>
<td>$P$-value</td>
<td>0.303333</td>
<td>0.353333</td>
</tr>
<tr>
<td>10% critical value</td>
<td>15.933561</td>
<td>15.873578</td>
</tr>
<tr>
<td>5% critical value</td>
<td>19.467372</td>
<td>18.545474</td>
</tr>
<tr>
<td>1% critical value</td>
<td>24.017914</td>
<td>23.780504</td>
</tr>
</tbody>
</table>

To determine the number of thresholds, the model is estimated by least squares allowing for zero, one, two and three thresholds. The test statistics $F_1$, $F_2$, and $F_3$, along with their bootstrap $p$-values are shown in column 1 in table 2. The test for a single threshold $F_1$ is highly significant with a bootstrap $p$-value of zero, and the test for a double threshold $F_2$ is also strongly significant with a bootstrap $p$-value of $2$.

$^2$ 300 bootstrap replications are used for each of the three bootstrap tests.
0.00333. On the other hand, the test for a triple threshold \( F_3 \) is not significant, with a bootstrap \( p \)-value of 0.30333. Thus, we conclude that there is evidence that there are two thresholds in the regression relationship. For the remainder of the analysis, we work with the double threshold model as follows

\[
\begin{align*}
\text{Growth}_{it} &= \mu_i + \beta_1 \text{Gini}_{it-1}I(\text{GDP}_{it-1} \leq \sigma_1) + \beta_2 \text{Gini}_{it-1}I(\sigma_1 < \text{GDP}_{it-1} \leq \sigma_2) \\
&\quad + \beta_3 \text{Gini}_{it-1}I(\sigma_2 < \text{GDP}_{it-1}) + \phi_i \text{Education}_{it-1} + \phi_2 \text{Distortions}_{it-1} + e_{it},
\end{align*}
\]

(3)

We also estimate another model, where we replace total educational attainment with male and female educational attainment as follows

\[
\begin{align*}
\text{Growth}_{it} &= \mu_i + \beta_1 \text{Gini}_{it-1}I(\text{GDP}_{it-1} \leq \sigma_1) + \beta_2 \text{Gini}_{it-1}I(\sigma_1 < \text{GDP}_{it-1} \leq \sigma_2) \\
&\quad + \beta_3 \text{Gini}_{it-1}I(\sigma_2 < \text{GDP}_{it-1}) + \phi_{i} \text{Male Education}_{it-1} \\
&\quad + \phi_{2} \text{Female Education}_{it-1} + \phi_2 \text{Distortions}_{it-1} + e_{it},\tag{4}
\end{align*}
\]

The variable \( \text{Male Education}_{it-1} \) is a measure of male educational attainment in country \( i \) in year \( t-1 \), and the variable \( \text{Female Education}_{it-1} \) is a measure of female educational attainment in country \( i \) in year \( t-1 \). The average years of education in the male and female population aged over 15 is extracted from Barro and Lee data on educational attainment. The test statistics \( F_1 \), \( F_2 \), and \( F_3 \), along with their bootstrap \( p \)-values are shown in column 2 in table 2. The results confirm the significance of two thresholds. The point estimates of the thresholds in both regressions are $1079.172759 and $1347.094572, respectively. More information can be learned from plots of the concentrated likelihood ratio function displayed in Figures 1-3. To examine the first-step likelihood ratio function which is computed when estimating a single threshold model, the first-step threshold estimate is the point where the likelihood function equals zero, which occurs at \( \sigma_1 = $1079.172759 \). There is a second dip in the likelihood ratio around the second-step estimate \( \sigma_2 = $1347.094572 \). Thus, the single threshold likelihood conveys information that suggests that there is a second threshold in the regression.

The regression slope estimates, conventional OLS standard errors, and white-correlated standard errors are reported in Table 3 for regression 1, and in Table 4 for regression 2. In the first regression, the estimates of primary interest are those on the Gini coefficient. Income inequality has a significant negative effect on economic growth with a coefficient of -0.004882, if real GDP per capita is below the first threshold $1079.172759. The coefficient is also significantly negative and equals -0.001222 if real GDP per capita is between the first and the second thresholds. On the other hand, the

\[3 \text{ 300 bootstrap replications are used for each of the three bootstrap tests.}\]
coefficient is not statistically significant, if real GDP per capita is above the second threshold $1347.094572$.

**Figure 1.** Confidence Interval Construction in Single Threshold Model.

**Figure 2.** Confidence Interval Construction in Double Threshold Model.
In the second regression, income inequality has a significant negative effect on economic growth with a coefficient of -0.004972, if real GDP per capita is below the first threshold $1079.172759. The coefficient is also significantly negative, -0.001325, if real GDP per capita is between the first and the second thresholds. On the other hand, the coefficient is not statistically significant, if real GDP per capita is above the second threshold $1347.094572.

These results suggest that there exists a threshold GDP per capita of around $1347, where income inequality hinders economic growth in countries with a lower GDP per capita, while income inequality does not impact economic growth in countries with a higher GDP per capita.
Table 4. Regression 2 Estimates

<table>
<thead>
<tr>
<th>Regressor</th>
<th>Coefficient Estimate</th>
<th>OLS SE</th>
<th>White SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male Education_{t-1}</td>
<td>-0.025437</td>
<td>0.011258</td>
<td>0.009483</td>
</tr>
<tr>
<td>Female Education_{t-1}</td>
<td>0.017721</td>
<td>0.011032</td>
<td>0.009244</td>
</tr>
<tr>
<td>Distortions_{t-1}</td>
<td>-0.000030</td>
<td>0.000012</td>
<td>0.000054</td>
</tr>
<tr>
<td>Ginii_{t-1}(GDP_{t-1} ≤ 1079.172759)</td>
<td>-0.004972</td>
<td>0.000724</td>
<td>0.002123</td>
</tr>
<tr>
<td>Ginii_{t-1}(1079.172759 &lt; GDP_{t-1} ≤ 1347.094572)</td>
<td>-0.001325</td>
<td>0.000610</td>
<td>0.000752</td>
</tr>
<tr>
<td>Ginii_{t-1}(1347.094572 &lt; GDP_{t-1})</td>
<td>0.000616</td>
<td>0.000546</td>
<td>0.000590</td>
</tr>
</tbody>
</table>

3. CONCLUSION

Theoretical proposals and empirical estimations have provided contradictory conclusions as to whether income inequality is propitious to economic performance, or whether it acts as an impediment to growth. Galor and Moav (2004) provide a reconciliation and argue that the replacement of physical capital accumulation by human capital accumulation as the prime engine for economic growth changes the impact of inequality on growth. In the early stage of development, inequality enhances the process of development by channeling resources towards those whose marginal propensity to save is higher, while in later stages, equality alleviates credit constraints on the investment in human capital and promotes economic growth.

This paper attempts to test empirically the finding that the effect of income inequality on economic growth depends on the development stage. A threshold estimation technique is utilized for 70 countries for the period between 1970 and 1999. The estimation suggests that there is a statistically significant threshold income per capita, below which the coefficient on the relationship between income inequality and economic growth is significantly negative and above which the estimate is not significant.

Data Appendix

The estimation uses annual data that covers the period from 1970 to 1999 for 70 countries, namely: Algeria, Australia, Austria, Bangladesh, Barbados, Belgium, Bolivia, Cameroon, Canada, Central Africa, Chile, Colombia, Cyprus, Denmark, Dominican Republic, Ecuador, El Salvador, Fiji, Finland, Germany, Ghana, Greece, Guatemala, Haiti, Hong Kong, Hungary, Iceland, India, Indonesia, Iran, Iraq, Ireland, Israel, Italy, Jamaica, Japan, Jordan, Kenya, Korea, Kuwait, Malawi, Malaysia, Mauritius, Mexico, Netherlands, New Zealand, Nicaragua, Norway, Pakistan, Panama, Papua New Guinea, Philippines, Poland, Portugal, Senegal, Singapore, South Africa, Spain, Swaziland,
Sweden, Syria, Taiwan, Tanzania, Tunisia, Turkey, United Kingdom, United States of America, Uruguay, Venezuela, and Zimbabwe. The variables used in the estimations are described in details as follows:

1. **Gini Coefficient**
   A detailed description of the Estimated Household Income Inequality dataset, compiled by the University of Texas Inequality Project, is provided in Galbraith and Kum (2004). This set combines data on the measure of dispersion of pay across industrial categories in the manufacturing sector, drawn from the industrial database published annually by the United Nations Industrial Development Organization UNIDO, to the information in the Deininger and Squire (1996) data, resulting in a dataset for the Gini coefficient referred to as the Estimated Household Income Inequality EHII.

2. **GDP Per Capita**
   The data for real Gross Domestic Product per capita (Laspeyres) is extracted from the Penn World Tables 6.2, which is obtained by adding up consumption, investment, government and exports, and subtracting imports in any given year, where the components are obtained by extrapolating the 1996 values in international dollars from the Geary aggregation using national growth rates. The growth rate of GDP per capita is given by the difference in the natural logarithm of the real GDP per capita in two consecutive years.

3. **Education**
   The data for total education, male and female education are derived from the Barro and Lee International Data on Educational Attainment in which they constructed estimates of educational attainment by sex for persons aged 15 and over. The values applied to several countries over five year intervals for 1970, 1975, 1980, 1985, 1990, 1995 and 1999. The estimation procedure began with census information on school attainment for males and females where the data came from individual governments as compiled by the UNESCO and other sources. We follow Forbes (2000) in using the average years of secondary schooling in the male population, and the average years of secondary schooling in the female population as proxies for male and female education, respectively. As the data is available only for the years 1970, 1975, 1980, 1985, 1990, 1995 and 1999, we use linear interpolation to derive the years-in-between.

4. **Distortions**
   Market distortions are proxied by the price level of investment extracted from the Penn World Tables 6.2. The variable is measured as the PPP of investment divided by the exchange rate relative to the United States times 100. The PPP of investment is the national currency value divided by the real value in international dollars. The PPP and the exchange rate are both expressed as national currency units per US Dollars. This measures how the cost of investment varies between each country and the United States.
It is meant to capture market distortions that affect the cost of investment such as tariffs, government regulations, corruption and the cost of foreign exchange.

REFERENCES
