

Income and inequality

Daniel Gandalf Hernandez

Instructor Christine Baker-Smith

QMSS

Master's thesis

Columbia University

Abstract

The developing world has on average seen higher economic growth rates during the last decades than the richer countries. If this trend continues within-country inequality will become the main driver of global inequality. This paper contains an overview of different theories of inequality and focuses on a previously hypothesized relationship between income and inequality. A two-stage model is created to show, using a newly available data set that on a country-by-country basis income per capita is not a good predictor of inequality, contradicting the theory. Additionally, economic shocks are studied and found to on average cause an increase in inequality, but that the increase is not statistically significant.

Introduction

From 1979 to 2007 the top one percent of wage earners in the USA saw their incomes grow by 156% while the bottom 90% only saw growth of 17% (Mishel and Sabadish 2012). CEO compensation grew from 18 times the average worker in 1965 to 210 in 2011 and inequality in the US has been steadily increasing in the last three decades. Brazil, however, long known as one of the most unequal countries in the world, has seen steady declines in inequality since 1998. Lopez-Calva and Lustig (2009) attribute 40% of the decline to a stronger focus on governmental transfer programs, such as increased social security benefits and the newly introduced Bolsa Familia program. Bolsa Familia provides benefits to poor families conditioned on children attending school and meeting certain other requirements.

Most countries have seen growth in income during the last decades, but the distribution of the gains varies markedly from country to country. For example, some of the East Asian Tigers¹ saw decreases in inequality as their economies expanded greatly from the 1960s onwards (Timmons 2010), while China has in the last 30 years seen impressive economic growth, but at the same time significantly increased inequality (Milanovic 2005). This paper contains an overview of different theories of what drives inequality, and focuses on a popular theory that country income (how rich a country is) has a direct relationship with how unequal the country is (Kuznets 1955). A counter-hypothesis that income is not a good predictor for inequality is proposed. Additionally, the overall effect economic shocks have on inequality in a country is studied. A recently compiled data set of inequality

¹ Hong Kong, Singapore, South Korea, and Taiwan

(Milanovic 2005) is paired with income per capita data from the World Bank for each available country from 1980 to 2007. The combined data set makes it possible to track income and inequality over time for each country, to ascertain if per capita income is a good predictor of inequality. A two stages analysis is used where each country is looked at in isolation to calculate its inequality trajectory as its income changes. The second phase combines the trajectories to see if they, on average, follow what Kuznets' theory postulates. It is my expectation that analysis will show that Kuznets' theory cannot be confirmed and that inequality does not have a simple direct relationship with income.

Economic growth and inequality

An early investigator of the relationship between inequality and country income outside of the classical, neo-classical, and Marxian views of surplus income accruing to landowners and owners of capital (Barber 1967) was Simon Kuznets (1955). Kuznets was interested in the relationship between the average monetary income gained by the population of country (income per capita) and how that income was distributed amongst the population. Kuznets considered inequality in the form of what percentage of total income was gained by fixed ordinal groups. For example (Kuznets 1955, p. 4) states that in the United States the top quintile (the 20% of the population with the highest income) earned 44% of the total income in the country during the 1940s². Kuznets suggested that economic inequality follows a pattern: as a country grows richer, inequality between different groups of people within that country will first rise and then fall. A country will

² More current research commonly uses the Gini coefficient as a measurement of inequality (Braun 1998, who also provides a discussion of different inequality measurements). A fuller discussion of Gini appears in a later section of this paper.

start with a majority of the population surviving by subsistence farming with limited income inequality (most people being poor). As the country starts to industrialize, some parts of the population move from agricultural production to higher paid industrial or service jobs in urban areas and inequality grows. The greater sophistication of the economy leads to increases in education and gains in power by the working class. The rising political power of the working class increases support for redistributive action, such as progressive taxation and welfare programs that lower inequality, as the poorer majority demand a share in the riches they see created around them. A graph with income on the x-axis and inequality on the y-axis would take on the shape of an inverted U-curve according to Kuznets (Figure 1).

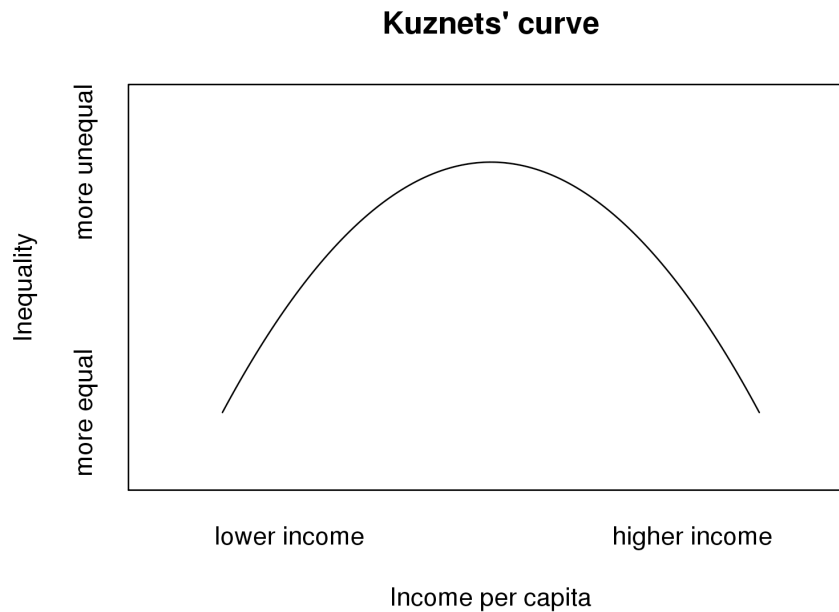


Figure 1 - Kuznets' curve

Kuznets' theory concerns the relationship between inequality and income within individual countries, and does not speculate on the differences seen between countries.

Inequality between countries have been explained by a multitude of historical, political and economic reasons, such as at which points technological innovations were introduced in the different countries and at what economic stage of the country democratization took hold (Barro 2000). Additional factors that influence inequality between countries are discussed in the background section of this paper.

Policy implications

In the political sphere the combination of Kuznets' ideas and the Laffer curve has inspired economic policy. The Laffer curve (Wanniski 1978) is a theory on the relationship between the level of taxation, overall economic activity, and total taxes raised. A very high tax rate removes workers incentives to work since they will not keep the fruits of their labor, and tax collections and economic activity are low. In the opposite case, with a low tax rate, incentives to work are high, since most gains accrue to the individual, and economic growth is at a higher level than in the high taxation case (Scully 2003). Because the level of effort by workers is tied to the level of taxation there exists an optimal tax rate that maximizes tax revenues. The Laffer curve takes its name by plotting the hypothesized total taxes raised as a function of the level of taxation. The plot will take on the shape of a concave curve, with no taxes raised at both a taxation rate of zero percent and at a hundred percent (since workers will not work if they cannot keep any gains in the latter case).

Advocates typically propose that taxes are too high, and that lower rates will stimulate economic activity plus increase tax revenues in the long run, as the economy grows faster under a new tax regime, as per the Laffer curve theory. The faster the economy grows, the faster the country travels along the Kuznets curve, leading to reductions in inequality, without explicit redistributive policy actions. The Kuznets theory

is used to justify the position that economic growth in itself will reduce inequality, while the Laffer curve is used to state that taxes need to be low to achieve fast economic growth – that using taxation to redistribute income to reduce inequality slows down the economic growth that would in itself reduce inequality.

Although these ideas are popular in certain political spheres there are serious misgivings among academics about their validity. For example H. W Arndt (1983) discusses what he refers to as the myth of “trickle-down economics”, and states, “it was not enough for development policy to pursue a high rate of growth of overall or per capita income. Policy needed to aim more directly at social objectives-employment, improved distribution, reduction of poverty” (p.1). Arndt contradicts the ideas that economic growth in itself will lead to lower inequality. Additionally, a fuller reading of Kuznets (1995) reveals an agreement with Arndt. The missing link in the political thinking is that Kuznets explicitly states that inequality will fall as the poorer majority gains additional powers and forces through redistributive actions, ideas also brought forth by Tam (2008). Some adherents of trickle-down economics seem to either ignore inequality as long as everybody sees some gains, or assume that inequality will drop even without the policies Kuznets mentions.

Background

Other factors of inequality

There exists significant literature on factors that influence inequality (Barro 2008; Tam 2008; Li and Zhou 2010). For example, demographic variables such as improvements in education, health, and urbanization have been shown to be effective in lowering inequality. Likewise, political variables including level of democratization and openness to trade ameliorate inequality, which are in line with Tam's (2008) findings of a political

Kuznets curve, touched upon in a subsequent section of this paper. Other idiosyncratic country characteristics such as cultural norms, country history, and natural resource endowments have also been shown to affect inequality and contribute to its country-by-country variation.

The causes of economic shocks and their policy responses are well discussed both in academia and in more widely read material, such as Joseph Stiglitz's "Globalization and Its Discontents" (2003). But there seems to be a dearth of studies on changes in inequality stemming from economic calamities, especially looking at patterns across crises and countries. For example, Rao (1999) examines the Asian Tigers³, together with other countries in the region that saw rapid growth, and inequality, but lacks data following the Asian Financial Crisis of 1997. Piketty's study of inequality in France (2003) finds that wage inequality has been extremely stable in the long run, and that the decline in income inequality seen during the first half of the 20th century was mostly driven by substantial shocks, such as the Great Depression and World War II. Kakwani and Pernia (2000) find that after the Asian crisis of 1997 there was a subsequent drop in inequality, but the study does not make comparisons with additional crises or regions.

An exploratory review done for this paper using inequality data provided by Milanovic (2005) shows that inequality increased significantly in Eastern Europe and Russia following the fall of communism. This is expected when moving from a system where nominally private for-profit enterprise did not exist and wages were set by bureaucrats, to a system of private enterprise and control of capital. The move to capitalism triggered immediate large drops in income during the early 1990s, with a follow

³ Hong Kong, Singapore, South Korea, and Taiwan

up crisis in the 1997, with contagion from economic problems in Asia (Svejnar 2002).

Studies mentioned earlier by Barro (2008), Banerjee and Duflo (2003), and Li and Zhou (2010) look at the relationship between income and inequality, and include various other variables, such as geographic region, but do not include variables specifically controlling for economic shocks.

James Galbraith at the University of Texas' Inequality Project presents some early findings on inequality and financial crises (1999) with data from 1996 and finds evidence for statistically significant changes, with general increases in inequality in poorer countries, especially in Latin America, but mixed changes for Europe and Asia. The study misses the period from 1997 to 2001 when economic shocks rolled across the globe, striking first Asia, before spreading to Russia and South America.

Country-level income explaining inequality

Kuznets' theory (1955) used average per capita income as an explanatory variable for inequality. At the first stages of development the population in a country is, to a large extent, poor and mostly agricultural. As industrialization begins, urbanization and gains in productivity in the city-bound industrial work force will increase inequality. Kuznets' examination of the existing data at the time, found a surprising drop in inequality as the economy kept growing. The prevailing thinking was that drivers of economic growth were inherently inegalitarian, with gains accruing at the top (Moran 2005). Kuznets' solution was to add institutional and political changes to the hypothesis. A growing and increasingly educated urban population would assert social and political power to force changes in legislation to bring in redistributive efforts and other labor friendly policies.

Korzeniewicz and Moran (2005) describe the initial positive reception to Kuznets' inverted U-Curve, but it came under increased attack in the 1980s and 1990s, when data became available showing rising inequality in richer countries. Some authors (Piketty 2006) have hypothesized that more advanced economies have entered into a new Kuznets' curve, driven by a new industrial revolution based on information technology, and that Kuznets' theory is a manifestation of innovations benefitting a select few at first, before diffusing into gains for the wider masses. Piketty additionally points out the lack of inequality data prior to the 1950s making longer time period studies difficult.

Robert J. Barro (2008) finds evidence for a Kuznets curve from the 1960s to the 2000s, investigating the shape of income inequality mapped against country income per capita. Barro uses an updated version of the Deininger and Squire dataset (1996), discussed in the data section of this paper, for Gini coefficients. He notes that while the curve exists, the effect is not pronounced when compared to the variance of inequality amongst countries at similar levels of income.

Li and Zhou (2010) likewise find evidence for an inverted curve for inequality over income, using Milanovic's Gini data set, with poorer countries experiencing increases in inequality as they grow, while richer countries have falling inequality. In addition, they detect a recent upturn in inequality in richer countries, which is in conflict with Kuznets' theory. The authors control for a variety of confounding factors and find that openness to trade, urbanization, and investment as a share of GDP are significant influences on inequality.

Democratization as a predictor for inequality

Henry Tam (2008), however, does not find a relationship between economic development and income inequality, but instead locates a political Kuznets curve using indices for democracy. The usual positive correlation between political and economic development explains why previously there have been observed relationships between income and economic development, with economic growth acting as a proxy for political growth. Tam outlines a possible explanation, echoing points made half a century earlier by Kuznets: initially an elite accumulates wealth and inequality increases with industrialization. As partial democracy starts taking hold, but with limited political power by the masses, a small, educated middle class emerges that gains clout to seize economical gains, and inequality increases further. The widening inequality leads to social unrest with pressure for further democratization. Suffrage increases and voting power reaches the poor masses, and they are able to assert themselves and force redistributive actions to lower inequality for their benefit. This study is very interesting in that it decouples income from inequality. If income is an inexact proxy for democratization it may explain why previous studies of income explaining inequality (as discussed earlier) and studies of inequality explaining income (as examined in the next section) have found contradictory results. Democratization may be a more fundamental driver of inequality than economic growth.

Studies by for example Rudra (2004) also find that higher democratization is associated with lower inequality. On the other hand, Timmons (2010) does not find a direct relationship between democracy and inequality, but points to theories about why increased democratization could lower inequality, in line with Tam's reasoning (Tam

2008). Democratization leads to changes in taxation benefiting the masses through greater redistribution of income and may also lead to greater access to public services, which previously was only available to more privileged groups, leading to improvements in education and health, which lessens poverty and decreases inequality. Timmons gives a few reasons for why his research failed to find a relationship between democratization and inequality, one of which is lack of data, another that gains in democratization is also associated with greater individual freedom, where now certain individuals are able to take advantage of new opportunities that were not available to them before. An example would be the oligarchs who were able to amass incredible wealth during the democratization and liberalization of markets in Russia in the 1990s (Guriev and Rachinsky, 2005).

Inequality explaining income

Kuznets (1955) ponders democratization and other political processes that are normally correlated with economic growth for falling inequality in more developed countries, but uses income as his predictor. Turning the relationship around, inequality has been used to explain economic growth in numerous studies. In the naïve case at the margins, total equality would depict a society with few economic incentives for additional productivity since any additional income one could gain would immediately be taken away and spread out amongst everybody else, to preserve universal equality. Conversely, in a totally unequal country, one person would gain all wages, and similarly, incentives for productivity in the rest of the population would be nil⁴.

⁴ This is the theoretical outcome of envisioning a society with either absolute equality or inequality. For example, an absolutely unequal society implies that everybody in the population save for one individual would have all their income taken away from them, without receiving anything in return, which is an absurd cases and would lead to immediate starvation for all save one..

Since efforts at changing levels of inequality involve redistributive policies (through, for example progressive income taxation), certain levels of inequality should lead to different levels of economic growth. "Too much" equality removes incentives for working hard, since you keep little extra of the results of your work, while "too little" equality leads to social and other issues that may be detrimental to growth, such as political unrest and large masses without access to education and health care. As seen by the different results in the following studies, there is little consensus on inequality's impact on economic growth.

Simpson (2009) examines available literature and data and argues that any efforts to reduce inequality is not only morally wrong, since it restricts voluntary trade, but will lead to lower economic activity, stunting income growth. Others, such as Barro (2008) find a negative relationship between inequality and economic growth (finding that higher inequality is associated with lower income growth), when looking at a large selection of countries. Banerjee and Duflo's analysis (2003) finds that any changes in inequality, positive or negative, lead to short-term reductions in economic growth. Forbes (2000) finds that increased levels of income inequality have a positive relationship with subsequent economic growth (that higher inequality is associated with higher income growth).

Income and inequality measurements

The level of (economic) development or 'richness' of a country is commonly measured using the gross domestic product (GDP). GDP, or national income, is the value of all goods and services produced and sold in a country for a given time period (Krugman and Obstfeld 2006). Gross national product (GNP) equals GDP with adjustments for income from abroad, and a few other factors. Krugman and Obstfeld state that, "The difference

between GNP and national income is by no means an insignificant amount, but macroeconomics has little to say about it, and it is of little importance for macroeconomic analysis" (p. 282). Numerous studies of inequality use GDP per capita to model country income, including Barro (2008), Sala-i-Martin (2006), and Schultz (1998). Heyse and Subarna (2006) mention various empirical studies on income and inequality using GDP measurements. GDP data will be used for country income in this paper, due to greater availability of GDP data, as well the common convention of using GDP for income. The two terms will be used interchangeably. Income includes all sources of capital, from normal employment and social welfare programs, to gains from illegal or unreported activities. An interesting side note is that GDP was first proposed by Simon Kuznets in 1937, when he was an economist at the National Bureau of Economic Research (Dickinson 2011).

Currency exchange rates and price levels complicate comparisons between countries, especially over time. Exchange rates can fluctuate significantly over short periods of time, driven, for example, by speculative capital movements (Firebaugh 2000), without significant underlying changes in the countries compared. Furthermore, different inflation rates complicate comparisons across time periods, where changing nominal values of currencies do not reflect actual buying power. A simple method to control for currency fluctuations and inflation is to translate each measurement of GDP to a common currency (usually the US dollar) at the prevailing exchange rate at the time of measurement, and then account for US inflation to bring the amount to the value of the dollar at a fixed point in time. This allows us to state that: country X_1 had an income per capita of Y_1 (at year 2000 USD values) in year Z and country X_2 had an income per capita of Y_2 that same year and if Y_2 is greater than Y_1 , then country X_2 was more economically

developed at that point in time. While the method above provides measurements in a common currency at a particular point in time, exchange rate biases are introduced in the analysis. A large part of goods and services within a country are not easily tradable across borders, and the general price levels of goods and services are not always reflected in the exchange rates (Firebaugh 2000). A simple example makes this problem more salient: while New York City and rural parts of the USA share the same currency (thus having an exchange rate of 1 to 1), price levels for housing, prevailing wages, and food vary wildly. Without controlling for these differences, and ignoring if a person lives in New York City or not, it is hard to make comparisons about how well off, or how much consumption is available to, that person at given levels of income.

In order to account for different price levels, thus mitigating the concern over exchange rates' inability to capture different prices by country, GDP is also reported in purchasing power parity (PPP) terms. GDP in PPP terms reported in fixed USD values accounts for price levels between countries, as well as inflation over time. Firebaugh (2000) states "Although early studies in economics used official exchange rates to convert local currencies to dollars, PPP-based estimation is now the industry standard. It is widely recognized that official exchange rates are badly flawed calibrators of currencies" (p. 333).

A key component not captured in GDP or country income is the distribution within the country. A country where all income in the country accrues to a single person will have the same GDP as a country where the same income is equally divided amongst all citizens. The distribution in income has significant impact on the characteristics of a country. Measurements of inequality seek to capture how income or wealth is distributed among

different individuals, and form a complement to GDP in describing the economic distribution in a country.

Measuring inequality based on income is an underestimation of the full level of inequality in a country due to the lack of accounting for accumulated assets. As an example, Wolff (2010) reports that in the USA 5% of the population earned 37% of all income (in 2006) but owned 62% of total wealth (in 2007). Even more striking, for the same years, the bottom 40% of the population earned 10% of total income, but their asset ownership only added up to 0.2%. Measuring inequality based on income or wealth is a choice between a flow and stock measurement. Income inequality is used in this paper since it more accurately reflects actual buying power of most individuals, simply due to fact that the majority has negligible stocks of wealth. In addition, GDP measures a flow of capital, which makes inequality based on income a more natural companion variable than the stock of individuals' wealth.

The most common measurement of inequality is the Gini coefficient, defined by Corrado Gini in 1912. The Gini coefficient is based on the Lorenz curve, which at each percentage point plots the proportion of income that is received by the population up to and including that point. The Gini coefficient is then derived from the area of the curve. Dorfman (1979) provides definitions and formulas for calculating Gini coefficients. A country where all citizens have the same income will have a Gini coefficient of zero while a country where one single individual captures all income will have a coefficient of one⁵.

⁵ It is common in studies to multiply the Gini coefficient by 100, to generate a number between 0 and 100, instead of zero and one. The data used and results presented in this paper have Gini coefficients measured from 0 to 100.

The Gini coefficient has certain drawbacks in trying to capture the concept of inequality. Braun (1998) mentions that the coefficient does not take into account different within-country price levels (as in the example with citizens in New York City and rural areas), and reacts more strongly to changes in incomes at the middle level compared to changes at either ends of the spectrum, giving it a tendency to remain fairly constant, even if underlying realities change significantly. Braun additionally describes other methods of measuring inequality that try to ameliorate some of the drawbacks with the Gini coefficient. For comparison, Braun uses data from the 1980 Census to calculate inequality estimates using eight different inequality methods⁶. He finds high intercorrelation between the values produced by the different methods, yet cautions that “the eight inequality measurements are not simply clones of one another” (p. 404). A lack of data, and familiarity with the Gini coefficient leads most inequality studies to still use it, or to use simpler measurements, such as percentages of income accruing to the top and bottom fifth of the population.

In-country and between-country inequality

Seventy percent of global inequality is between-nation inequality, because of the vast differences in GDP per capita between the richest and the poorest countries (Firebaugh 2000). Economic growth rates have, in the past two decades, either been similar or higher for poorer countries compared to richer ones. The, on average, higher growth rates for poorer countries have allowed them to start catching up with the richer ones. Sala-i-Martin (2006) highlights gains made especially in Asia in the last two decades

⁶ The different inequality measurements methods are: Gini, Coefficient of Variation, Theil index, Nelson ratio, and four variations of the Atkinson measure (Braun 1998)

of the 20th century that significantly lowered global inequality. During this time, many fast growing developing countries have simultaneously seen increased within-country inequality. The overall effect has been a trend towards global equality on a country basis, but more within-country inequality. If the trend continues and developing countries keep posting higher average growth rates, within-country inequality will in the future become the main factor in global inequality, further increasing the topic's importance.

Hypothesis

While there is little consensus regarding the Kuznets inverted U-curve (Moran 2005), research that finds a relationship between inequality and income usually yield only a small explanatory effect. Calculations by, for example, Robert Barro (2008) show that country income can explain levels of inequality, but Barro remarks, "this curve does not explain the bulk of the observed variation in income inequality across countries or over time" (p. 8). Large variations in inequality exist between countries, as discussed earlier. As another example, countries in Eastern Europe find themselves at much lower levels of inequality than Latin American countries, even when sharing similar levels of income (Milanovic 2005). Countries that have moved from lower to higher levels of income during the last thirty years had on average lower levels of inequality at the outset of the available data, compared to countries at similar lower income levels today (Milanovic 2005). Additionally, countries such as the USA experienced increased inequality since the 1980s, while Japan saw drops, and Sweden went through little overall change during the same period (Milanovic 2005). I propose that income per capita (the independent variable) is not a strong predictor of inequality (the dependent variable), as proposed by Kuznets. When

looking at individual countries I expect to not see a significant pattern of richer countries on average being further along their Kuznets curves.

A secondary hypothesis is that economic crises will more adversely affect the less well off, leading to increases in inequality, in line with some of the results seen by Galbraith (1999).

Data

Reliable data availability has been an issue in the study of income inequality. Deininger and Squire tried to address this in 1996 with the creation of a comprehensive data set comprising 108 countries, made public by the World Bank (Deininger and Squire 1996). The researchers also enumerate various complexities encountered with data accuracy and availability. For example, data can be measured based on households or individuals, based on income or calculated from consumption levels. Data also excludes nonmonetary compensation, which in some countries is a significant part of total income. Additionally Deininger and Squire had to discard data based on lack of representativeness of the samples used for certain countries.

Branko Milanovic (2005) at the World Bank has, between 2004 and 2010, gathered data from five different inequality data sets (including Deininger and Squire's) and periodically made the aggregate set available as the "All the Ginis database" covering the period from 1950 to 2007. In addition Milanovic marks each data point if it is based on income or consumption, net or gross base, or on individual or household basis, which makes comparisons across countries and time periods easier, by only focusing on data points with the same underlying characteristics. The author further echoes the comments by Deininger and Squire on the difficulty in retrieving reliable inequality data.

The previous absence of robust Gini data, combined with ongoing questions of data quality may be a partial explanation for the different conclusions drawn about inequality and economic growth, as reviewed earlier. Reiterating, Heyse and Subarna (2006) find that an increase in inequality in poorer countries is associated with increased economic growth, while Banerjee and Duflo (2003) notices that any change in inequality is associated with a decrease in economic growth rates. Both studies used Gini coefficients from the Deininger and Squire 1996 data set.

The situation has improved in the last few years with the help of Milanovic and the maintainers of the underlying datasets upon which he draws, but significant holes still exist for periods of time for which no data is available, especially for developing countries, where household surveys of income are often lacking. Data also exists that eludes simple explanations. For example, Norway had a Gini coefficient of 25.7 in 1995 that jumped to 32.4 in 1996, during a period without significant political or economical events in the country or in Europe as a whole.

Data on GDP on a PPP basis is available from the World Bank in their "World Development Indicators" (WDI) database. Numbers are given for a large majority of countries from 1980 until 2009. It would be advantageous to have a longer time series, since the All the Ginis database has sporadic data reaching back to the 1950s. Unadjusted GDP data is available for the longer time frame, but for reasons described earlier of the importance of using price-adjusted levels, the analysis presented is done using the available PPP data.

The data used was constructed from the Milanovic Gini and the WDI income sets. Each data point in the Milanovic set that had income measured based on income (as

opposed to consumption) on an individual level was considered. For each Gini point the corresponding income value was located in the WDI data set, and joined to create one observational pair of a Gini coefficient and income per person on a PPP level for a particular country in a given year. The data set used for analysis spans 1980 until 2007 and consists of a total of 765 samples. The WDI data set has very good coverage, and most missing pairs are due to lack of Gini data. Table 1 presents descriptive statistics of the joining of the Milanovic's All the Ginis and the WDI datasets, by country and year, both aggregate and broken down by region. Note that all income numbers are in logarithms.

Table 1 - Descriptive statistics of Gini and GDP data

Variable	Mean	S.D.	Min	Max
World Gini	40.24	12.40	17.52	75.40
Africa	45.88	10.51	25.90	63.03
Asia, Middle East	38.58	6.612	23.38	50.80
Eastern Europe	29.40	7.196	17.52	62.14
Latin America	54.42	6.256	38.93	75.40
West, AUS, NZL, TUR	32.49	4.557	22.44	47.20
World log.GDP	9.143	0.912	6.261	11.09
Africa	7.390	0.876	6.274	8.984
Asia, Middle East	8.467	1.339	6.261	10.54
Eastern Europe	8.980	0.496	7.354	9.954
Latin America	8.736	0.432	7.008	9.664
West, AUS, NZL, TUR	10.128	0.272	8.901	11.09
	Samples	Countries	Missing samples	% Available
World	765	105	2175	35%
Africa	23	19	509	5%
Asia, Middle East	76	12	260	29%
Eastern Europe	176	26	552	32%
Latin America	260	24	412	63%
West, AUS, NZL, TUR	230	24	442	52%
Year range	1980-2007			

Lack of data presents a problem, especially in Africa, where only 5% of possible data points are available. Asia is also underrepresented. Comprehensive data exists for China,

Japan, and South Korea, amongst others, while unfortunately India is fully absent. All available points were used for the analysis in this paper. The underrepresentation of many poorer countries from the dataset (especially in Africa) will lead to less explanatory power for poorer countries, but since some of the major growing countries are present with good data (such as China), results will be available for countries representing the full spectrum of the income scale. Shortage of data will make validating a relationship harder. Conversely, if a relationship is not found, lack of data complicates deciding if the relationship is missing simply because of lack of data, or because Kuznets' theory is not accurate (the main hypothesis in this paper).

The combined dataset was mined for economic shocks, where a shock was defined as a drop of at least 2.5% in GDP from one year to another. Table 2 lists the shocks, grouped by region. The availability of data in this set is skewed towards Eastern Europe and Latin America, which will hurt the generalizability of any results.

Table 2 - Frequencies of economic shocks

	Shocks	Samples available
World Sample size	115	1102
Africa	5	12
Asia, Middle East	4	140
Eastern Europe	58	192
Latin America	40	349
West, AUS, NZL, TUR	6	409
Countries in sample		105

A large part of the shocks stem from Eastern Europe's dismantling of communism and tumultuous road to capitalism (for an overview, see Svejnar 2002). African countries have experienced significant economic shocks, but due to lack of African Gini data, very few shocks from the region can be studied.

Methodology

The analysis is broken down into three parts, with the first two looking at income's relationship with inequality using increasingly disaggregated parts of the data, and the third examining the impacts of shocks on inequality. The initial method fits an ordinary least squares (OLS) linear regression on the data. Income is used both as a linear and as a squared term to generate a fitted curve. Heyse and Subarna (2006) use OLS regressions to study how inequality affects GDP growth. Barro (2008) shows an example of using OLS linear regression to estimate a Kuznets curve using GDP per capita to estimate Gini coefficients. The goal of the first simple model is only to show that a Kuznets curve can be constructed from the data, when using all available data and discarding problems of dependence between data points for a given country across time, and shortage of data for individual countries.

The second method dives into the individual country level. An OLS linear regression model is run separately for each country, using income to predict inequality, without the squared term of the initial model to generate a single coefficient for each country. Coefficients from the individual regressions are then used as the dependent variable for a second OLS linear regression, where the independent variable is the average income of the country the coefficient belongs to. If a Kuznets curve exists, countries that are poorer should, on average, show higher coefficients than richer ones, as will be illustrated in the results section. This model can be thought of as using the roughly 30 years worth of data (from the combined Milanovic and WDI data set spanning 1980 to 2007 constructed above) to estimate a tangent for each country on where it is on its hypothetical Kuznets curve. Richer countries should have lower tangents, since they should be further ahead on their

curve (since the tangent of a point on an inverted-U/concave shaped curve will decrease as the point moves to the right along the curve). Using a regular OLS regression for the two-stage model is appropriate, because we have collapsed each country's data into a single coefficient, removing the time-series aspect of the data (that points belonging to a particular country are not independent across time, suffering from serial correlation, Wooldridge, 2009).

The last analysis, which is separate from the first two, is an attempt to understand movements in inequality and examines if there is a change in inequality, as defined by the difference in average inequality during the five years before and after an economic shock. The method used is a simple analysis to see if the change in inequality after a shock is statistically different from zero.

Results

Initial method

Using all available data in aggregate shows that there is a statistically significant curve tracing through the data, in line with Kuznets' hypothesis (Table 3). Figure 2 shows all available Gini/GDP pairs, and the fitted curve from the model. The drawback of this aggregate model is that it does not take into account that certain data points are related, coming from the same country, as mentioned above. Additionally, for some countries there are very few data points, so little or nothing is known about their changes in inequality as their income changes. Such countries still contribute to the model above, even though we cannot say if they are individually following the expected growth path hypothesized by Kuznets. For these reasons a more detailed model is warranted, especially since the

Kuznets' theory is about the path an individual country takes, and not on what the data together for all countries shows.

Table 3 - Global model

logged GDP	48.3690
	(8.079)***
(logged GDP) ²	-3.044
	(-9.022)***
Intercept	-145.01
	(-5.491)***
Observations	765
Adjusted R ²	0.24

*** = significant at 0.1%

Effect of GDP on Gini coefficient

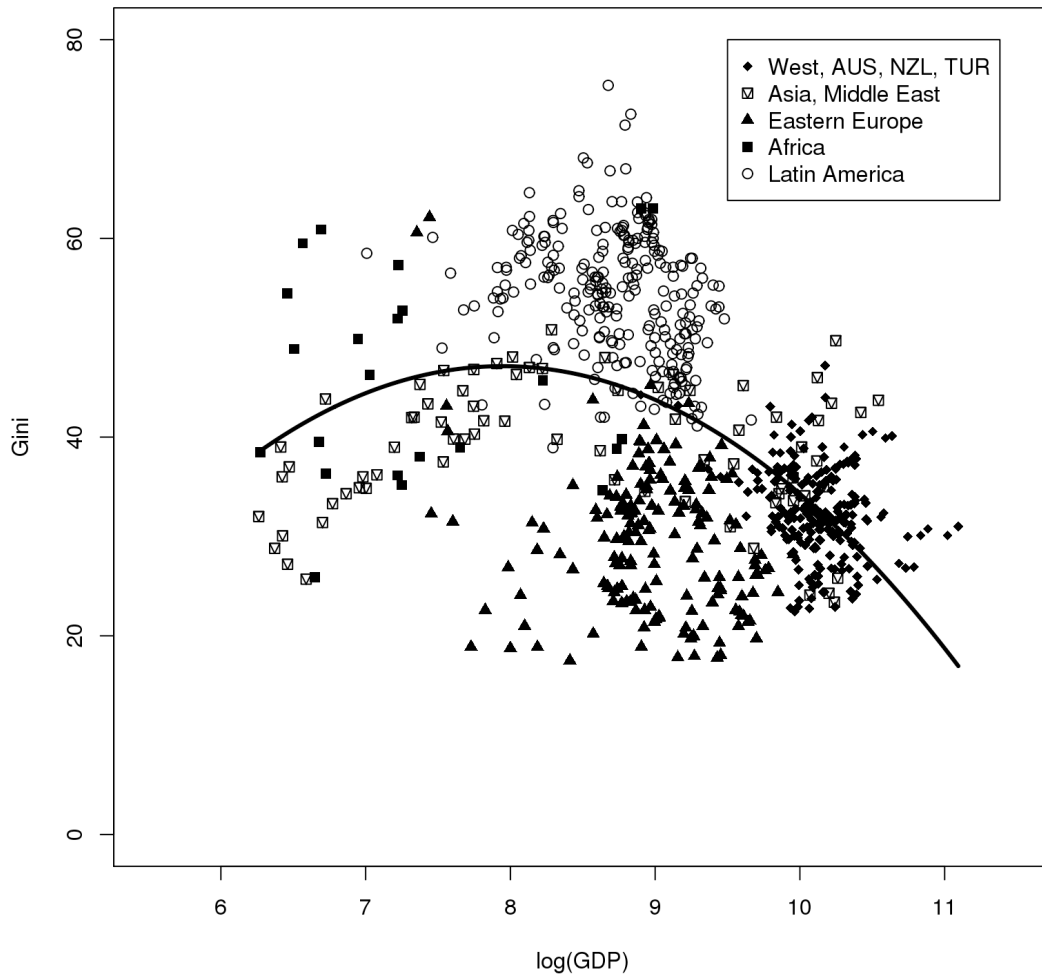


Figure 2 - Effect of GDP on Gini coefficient

Individual country method

The second model runs an individual regression with income explaining inequality for each available country, where there were at least two data points available (a total of 72 countries). The average income for each country was then used in a regression to explain the coefficient. If the Kuznets theory holds on an individual country level, countries that were on average richer during the period should have lower coefficients/tangents (being further along the Kuznet's curve) than poorer ones. Figure 3 and Table 4 show the coefficients for the individual countries. The left side of Figure 3 contains only coefficients that are significant at the 5% level, while the right side includes all coefficients. Visually, we would expect a downward sloping trend in both graphs. In contrast, the results show upward sloping (but not statistically significant) trends, contradicting the Kuznets theory.

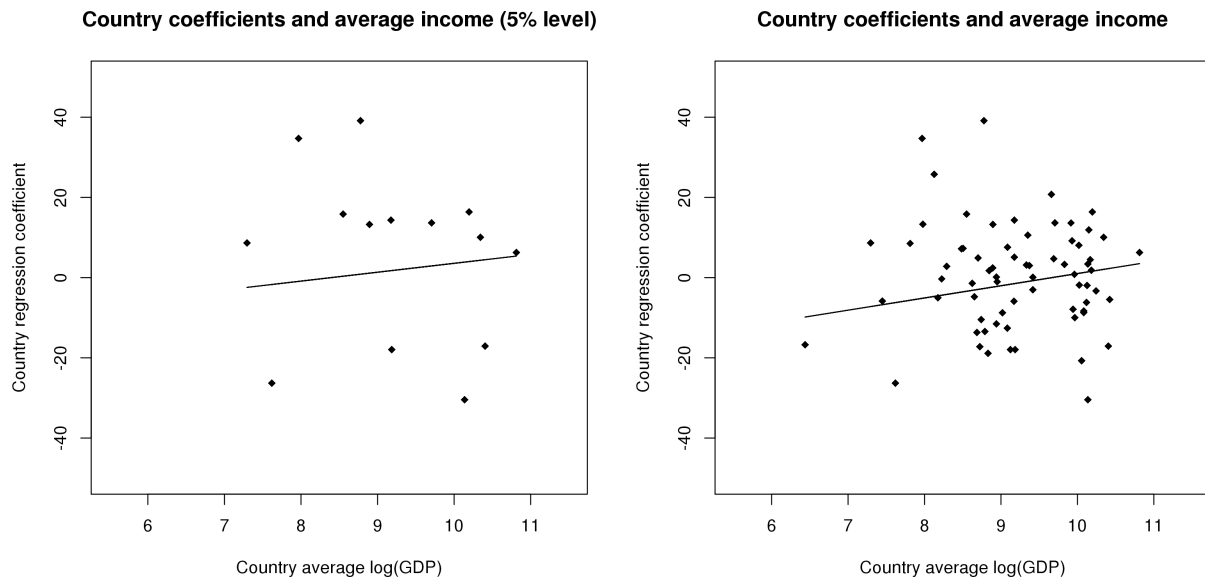


Figure 3 - Country coefficients

Table 4 - Country coefficients and average income

Variable	Mean	S.D.	Min	Max
Country coefficient	-1.540	17.06	-76.03	39.13
Average income	9.161	0.906	6.436	10.81
Year range				1980-2007
Sample size				71

The average income of each country was used in an OLS linear regression to predict the country regression coefficients. Two versions were run, one including only significant coefficients (5% level), and one using all coefficients. No significant results were found, as seen in Table 5. For robustness, the minimum, maximum, and median income of each country was also used instead of the average income as input with no change in the significance of the results.

Table 5 - Country level model

Coefficients in sample	All	0.05 level
Mean logged GDP	3.044 (1.360)	2.218 (0.354)
Intercept	-29.421 (-1.428)	-18.611 (-0.324)
Observations	71	15
Adjusted R ²	0.01	-0.07

There is no statistically significant relationship on the country level between income and inequality using the available data when we take into account what we know about each individual country, as opposed to using all data in aggregate, as in the initial method.

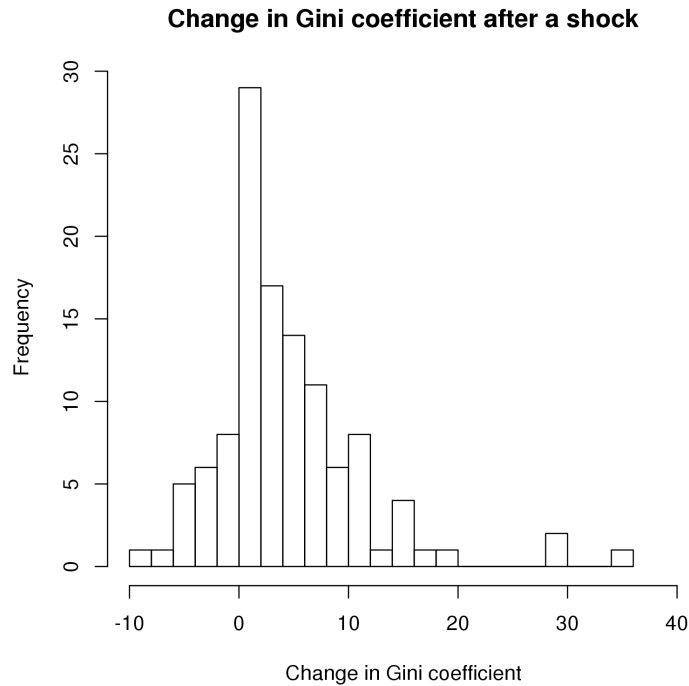
Economic shocks impact on inequality

Lastly, economic shocks were studied for what impacts they have on inequality. A drop in at least 2.5% of GDP was on average associated with an increase in inequality. The

data was largely driven by multiple crises in Eastern Europe during the 1990s, but the effect was still seen for other countries. Unfortunately, due to large standard deviations no statistically significant conclusions can be drawn. The results shown here mirror the different results seen, and lack of consensus, in studies of economic shocks discussed earlier, where Kakwani and Pernia (200) saw drops in inequality after a crisis, while Galbraight (1999) saw increases in inequality or mixed results. It is clear that better data is needed for follow up studies on the impact of economic shocks. Table 6 summarizes the results, and Figure 4 provides a histogram of the change in inequality after each economic shock, again showing that the majority of shocks saw increased inequality within the affected country. As seen in Figure 4, the most common occurrence in a shock was a slight increase in the Gini coefficient (increased inequality), but with high enough variability to not be able to draw any firm conclusions. For robustness, the definition of a shock was also set at values of a drop in GDP of 2%, 3%, 4% and 5%, without significant change in the results with using a 2.5% drop as cutoff.

Table 6 - Effects on inequality from economic shocks

Region	World	World excluding Eastern Europe	Eastern Europe
Mean change in inequality	4.38	0.60	8.02
Standard deviation	6.70	2.61	7.42

**Figure 4 - Change in inequality after an economic shock**

Conclusions

This paper set out to study two topics. Primarily, to investigate the relationship between income and inequality using the latest available data, to hopefully shed a small amount of new light on Kuznets' U-curve theory. The secondary topic was to investigate what the typical effect of an economic shock is on inequality. The study of economic shocks revealed that that, on average, countries see increased inequality afterwards, but the results were not statistically significant.

Returning to the main topic, using all data in aggregate a Kuznets' curve appears, but when each individual country's path is taken into account, the effect disappears. There are

multiple possible reasons for this. One is that the Kuznets' curve may just be an artifact of the aggregate data. As an example, revisiting Figure 2, we can see that Latin America clusters in the middle spectrum of income and shows consistently high inequality, helping to give an overall curved shape to the data. The high inequality seen in that middle-income region may be more due to the societal and economic history of the region than its place on a hypothesized curve.

Another reason may be that there is simple not high quality data (as discussed by Milanovic 2005, and noted in numerous other sources mentioned in this paper) to do a strong country-by-country model of the data, thus hiding a potential Kuznets curve, as well as exemplified by the ongoing discussions and multitude of papers revisiting Kuznets' hypothesis more than half a century after its formulation.

One appealing theory mentioned earlier was put fourth by Henry Tam (2008). Tam proposes that level of democratization drives inequality, and that income may be an inexact proxy variable for how democratic a country is, since richer countries on average score higher in indices created to measure democratization.

By using a country-by-country breakdown to study the effects of income on inequality I showed that, with the currently available data, income is not a good predictor of inequality, contradicting Kuznets' hypothesis. It is my hope that these results can in some small way help to either put to rest ideas that economic growth is the only thing needed to combat inequality, or at least to show that there is need for both better data and more studies before similar claims can be made and advocated as political policy.

Lower income countries have on average seen faster economic growth rates than richer countries, exemplified by the tremendous economic advances in China and India

during the last two decades, but also seen in high growth in poor countries in Africa, Latin America, and other regions. At the same time inequality in countries like China has increased significantly. If this trend continues between-country-inequality will continue to decrease in the coming decades, as poorer countries start catching up to richer ones. In-country inequality will become the main driver of global inequality. The income of a person will be less explained by which country she resides in, but by where she is located on the socioeconomic scale of that country, lending further importance to the study of models that look at each country on an individual basis, as well as to the importance of well contemplated political decisions on how to approach inequality.

References

- Arndt, H.W. (1983). The "Trickle-down" Myth. *Economic Development and Cultural Change*. Vol 32, No 1, pp. 1-10. Retrieved 10 Oct 2010 from JSTOR
- Banerjee, Abhijit V and Duflo, Ester. (2003). Inequality and Growth: What can the Data Say? *Journal of Economic Growth*. Vol 8. No 3. pp. 267-299
- Barber, William J. (1967). *A History of Economic Thought*, First Wesleyan Edition 2009, Wesleyan University Press, pp.
- Braun, Denny. (1988). Multiple Measurements of U.S. Income Inequality. *The Review of Economics and Statistics*. Vol 70. No 3, pp. 398-405. Retrieved 20 Nov 2010 from JSTOR
- Deiningner, Klaus, and Squire, Lyn. (1996). A New Data Set Measuring Income Inequality. *The World Bank Economic Review*. Vol 10, No. 3, pp. 565-591. Retrieved 21 Nov from JSTOR
- Dickinson, Elizabeth. (2011). GDP: a brief history – One stat to rule them all – *Foreign Policy*. January/February 2012. Retrieved 16 Sep 2012 from http://www.foreignpolicy.com/articles/2011/01/02/gdp_a_brief_history
- Dorfman, Robert. (1979). A Formula for the Gini Coefficient. *The Review of Economics and Statistics*, Vol 61. No 1, pp. 146-149, Retrieved 16 Nov 2010 from JSTOR
- Firebaugh, Glenn. (2000). The Trend in Between-Nation Income Inequality. *Annual Review of Sociology*. Vol 26, pp. 323-339. Retrieved Nov 16 2010 from JSTOR
- Guriev, Sergei and Rachinsky, Andrei. (2005). The Role of Oligarchs in Russian Capitalism. *The Journal of Economic Perspectives*. Vol 19, No. 1 (winter, 2005), pp. 131-150. Retrieved Sep 15 2012 from JSTOR
- Kakwani, Nanak and Pernia, Ernesto M. What is Pro-poor Growth? *Asian Development Review*. Vol 16. No 1. pp. 1-22. Retrieved 21 Nov 2010 from JSTOR
- Korzeniewicz, Roberto Patricio, and Moran, Timothy Patrick (2005) *Theorizing the Relationship between Inequality and Economic Growth*. *Theory and Society*. Vol 34. No. 3, pp. 277-316. Retrieved 16 Nov from JSTOR
- Krugman, Paul, and Obstfeld, Maurice. (2006). *International Economics Theory & Policy* (7th ed). Pearson, Addison-Wesley

- Li, Kui-Wai, and Zhou, Xianbo. (2010). A nonparametric and Semiparametric Analysis on the Inequality-Development Relationship. Presented at the the WTO Geneva Trade and Development Workshop. Retrieved 30 Oct 2010 from http://www.wto.org/english/res_e/reser_e/gtdw_e/wkshop10_e/autumn10_e.htm
- Lopez-Calva, Luis F, and Lustig, Nora (2009) .The recent decline of inequality in Latin America: Argentina, Brazil, Mexico and Pero. Ecineq, Society for the Study of Economic Inequality. Working Paper Series, ECINEQ 2009-140, October 2009. Retrieved 7 Aug 2012 from <http://www.ecineq.org/milano/WP/ECINEQ2009-140.pdf>
- Milanovic, Branko. (2005). Worlds Apart: Measuring International and Global Inequality. Princeton: Princeton University Press. Data available at <http://econ.worldbank.org/projects/inequality>. Retrieved 21 Nov 2010
- Mishel, Lawrence and Sabadish, Natalie. (2012). CEO Pay and the Top 1% - How executive compensation and financial-sector pay have fueled income inequality. Economic Policy Institute. Issue Brief #311, May 2, 2012. Retrieved 7 Aug 2012 from <http://www.epi.org/files/2012/ib331-ceo-pay-top-1-percent.pdf>
- Moran, Timothy Patrick (2005) Kuznets's Inverted U-Curve Hypothesis: The Rise, Demise and Continued Relevance of a Socioeconomic Law. Sociological Forum. Vol 20, No 2, pp. 209-244. Retrieved 10 Oct 2010 from JSTOR
- Piketty, Thomas. (2003). Income Inequality in France, 1901-1999. The Journal of Political Economy. Vol. 111. No 5. pp. 1004-1042. Retrieved 16 Nov 2010 from JSTOR
- Piketty, Thomas. (2006). The Kuznets Curve: Yesterday and Tomorrow. In Mookherjee, Dilip (Ed.), Banerjee, Abhijit Vinayak (Ed.), Benabou, Roland (Ed.) Understanding Poverty. pp. 63-72 Oxford University Press
- Rudra, Nita (2004). Welfare Spending, and Inequality in the Developing World. International Studies Quarterly, Vol. 48, No. 3 (Sep., 2004), pp. 683-709. Retrieved 15 Sep 2012 from JSTOR
- Sala-i-Martin, Xavier. (2006). The World Distribution of Income: Falling Poverty and Convergence. Period. Quarterly Journal of Economics, Vol 121. No 2. pp. 351-397. Retrieved 22 Nov 2010 from <http://www.mitpressjournals.org/loi/qjec>

Scully, Gerald W. (2003) Optimal Taxation, Economic Growth and Income Inequality. Public Choice. Vol 115.

No 3/4 (Jun., 2003) pp. 299-312. Retrieved Sep 14 from JSTOR

Schultz, Paul T (1998) Inequality in the Distribution of Personal Income in the World: How It Is Changing and

Why. Journal of Population Economics. Vol 11. No. 3. Pp. 307-344. Retrieved 12 Dec 2010 from JSTOR

Stiglitz, Joseph E. (2003) Globalization and Its Discontents. Norton and Company

Subarna, Samantha K, Heyse, Allison (2006). Income Inequality and Economic Growth in Developing

Countries: An Empirical Analysis. Indian Journal of Economics and Business. Dec, 5(2), pp.243-254.

Retrieved 12 Dec from http://findarticles.com/p/articles/mi_m1TSD/is_2_5/ai_n25012649/

Svejnar, Jan. (2002). 'Transition Economies: Performance and Challenges', Journal of Economic Perspectives.

Vol 16, Number 1, Winter 2002, pp. 3-28. Retrieved 22 Jul 2012 from JSTOR

Tam, Henry. (2008). 'An economic or political Kuznets curve?', Public Choice. Vol 134, No 3/4. 367- 389.

Retrieved 16 Nov 2010 from JSTOR

Wanniski, Jude. (1978). Taxes, Revenues and the 'Laffer Curve'. The Public Interest. Winter 1978, pp. 4-5.

Retrieved 10 Oct 2010 from http://www.nationalaffairs.com/public_interest/detail/taxes-revenues-and-the-laffer-curve

Wolff, Edward N. (2010). Recent Trends in Household Wealth in the United States: Rising Debt and the

Middle-Class Squeeze – an Update to 2007. Working Paper No. 589. Levy Economics Institute of Bard

College. Retrieved 21 Nov 2010 from <http://www.levyinstitute.org/publications/?docid=1235>

Wooldgridge,, Jeffrey M. (2009). 'Introductory Econometrics – A Modern Approach" 4th Edition. South-Western

Cengage Learning