5-2012

The Human Opportunity Index and Income Inequality in Brazil

Joseph Van Matre
University of Arkansas, Fayetteville

The Human Opportunity Index and Income Inequality in Brazil

by

Joseph Charles Van Matre

Advisor: Dr. Andrew Horowitz

An Honors Thesis in partial fulfillment of the requirements for the degree Bachelor of Science in International Business.

Sam M. Walton College of Business
University of Arkansas
Fayetteville, Arkansas

May 11, 2012

Abstract

This study calculates the Human Opportunity Index (HOI) proposed by de Barros et al. (2009) and a Gini Index of household income using nationally representative data from Brazil for the years 2001 to 2008. Macroregional HOIs and Gini Indexes were also calculated for the years 2001, 2005, and 2008. No statistically significant, linear trends were found over the eight year period for national HOIs. A statistically significant, linear, downward trend was found for the Gini Index of household incomes. Also, there was a statistically significant decrease in the macroregional variance of HOI scores, indicating a convergence of opportunity equity among macroregions. No such convergence was found for the macroregional Gini Indexes.

JEL classification:
D31, D63

Keywords:
Income Inequality, Equality of Opportunities, Brazil

Introduction

Outcomes in income distribution, access to government services, or general quality of life can be determined by many factors. The distributions of these are rarely egalitarian and often disproportionally benefit the better-off in society. Game theoretical and behavioral studies, such as Hoffman, McCabe, and Smith (1996) and Falk, Fehr, and Fischbacher (2003), show, however, that individuals have a preference for equality, many times even when it is contrary to rational self-interest. de Barros et al. (2009) conclude that the literature "reveals a remarkably widespread "taste for fairness" (p 46).

In addition to an intrinsic human desire for equity, Galor and Zeira (1993), Banerjee and Newman (1993), and Aghion and Bolton (1997), among others, have shown that the distribution of wealth can affect total societal output. Imperfect credit and labor markets mean that educational, investment, or entrepreneurial opportunities can be misallocated relative to the socially optimal distribution. Ferreira and Walton (2006) use the example of an unintelligent student being sent to the best schools because of his or her family's ability to pay, while a bright student from a low-income family may have to drop out. Efficient credit markets would alleviate the bright student's family's inability to finance his or her education while inefficient markets rob him or her of that opportunity and decrease total social output.

Goldstein and Udry (2008) offer an example in Ghana where inequality of opportunity robs society of increased total output. In the area studied, land is allotted based on societal convention, rather than on western conventions of property rights. Thus, property rights are generally based on societal hierarchy, and property security is based on an individual's standing within that hierarchy. The authors found that individuals that were not within the power structure of the societal hierarchy were more likely to have their land confiscated if they left their land unsewn for a time (in order to recover vital nutrients in the soil). Women, in particular, were less likely to do so. Consequently, a woman's land was generally less productive than a man's. This lack of productivity is representative of the societal loss that is experienced due to inequality of opportunity.

Inequality of outcomes in all of these areas is difficult to deal with conceptually as there is a major divide in the mechanisms by which to alleviate inequality. Should inequality be alleviated when it is observed through resource transfer programs, or should the inherent sources of inequality be alleviated such that observable inequality is reduced? The latter objective is formalized by Sen (1985), Rawls (1971), and Roemer (1998), among others. The thrust of this literature is to examine "ex-ante potentials" as opposed to "ex-post realizations"; that is, to equalize individuals' opportunities, not necessarily their outcomes. Sen (1985) discusses equalizing *capabilities*, Rawls (1971) focuses on allocating "primary goods" to the least advantaged groups, and Roemer (1998) formalizes the idea of inequality of opportunity. This ideological shift from *what will be* to *what could be* underpins the idea of equality of opportunity. The Roemerian concept of inequality of opportunity forms the conceptual base of the present research.

This study aims to investigate the relationship, if one exists, between a relatively new metric or inequality of opportunity, the Human Opportunity Index proposed by de Barros et al. (2009), and the Gini Index, a measurement of income inequality. Correlation coefficients and OLS estimation will be used to determine if a linear relationship exists between the two measurements at various time intervals.

Roemer's Concept of Equality of Opportunity

Roemer (1998) formalizes the concept of equality of opportunity first by distinguishing between two types of opportunity equity. He describes these two types as "leveling the playing field" and the "nondiscrimination principle". The nondiscrimination principle is most connected to the idea of ex-post realizations and refers the idea that an individual's gender, family background, or other personal attributes should not be used as a basis for allocating opportunities or resources if such attributes are not relevant to a particular resource. Not using race as a discriminating factor in job placement is an example of this principal. Leveling the playing field is more closely related to ex-ante potentials and is described by Roemer as a continuation of the nondiscrimination principle. It requires not only that personal characteristics unrelated to a job, for example, not be considered in job selection, but also that disadvantages that an individual may have because of his or her gender, family background or other personal attributes be corrected for as well. The latter is the concept referred to when this paper discusses inequality of opportunity.

Roemer goes on to formalize the concept of equality of opportunity in this sense. He does so by first decomposing the determining factors of an individual's outcomes into two categories which he calls "circumstances" and "efforts". Roemer defines circumstances as factors that are exogenous to an individual, such as race, gender, place of birth, or family structure. Efforts, by contrast, are factors that are endogenous to the individual, such as motivation or personal interests. He proposes that an individual's set of circumstances should not play a role, directly or indirectly, in outcomes.

Roemer proposes that equality of opportunity is a state when an individual's outcomes are based on efforts and are not directly or indirectly affected by his or her circumstances. In order to measure this kind of inequality, he suggests that a population of interest be broken up into groups based on identifiable circumstances, such that each group includes only individuals homogeneous in their circumstances. Roemer asserts that equality of opportunity is achieved when the measured outcome of the different homogeneous groups is equal and thus the heterogeneity of outcomes would only be observed within a homogeneous circumstance group based on differences of "effort" put forth.

Consider, for example, a classroom of 20 students in a particular course. For the sake of simplicity, assume that the only differentiating circumstances are the student's gender and whether or not a particular student comes from a one or two-parent household. Further assume that the only measurable outcome of interest is whether the student passes that particular class. Using a Roemerian sense of equality of opportunity, the class should be divided into four homogeneous circumstance groups: males from two-parent households, males from single-parent households, females from two-parent households, and females from single-parent households. For there to be equal opportunity under Roemer's system, the pass rates of the homogeneous groups should be equal. That is, simply knowing the circumstances of an individual should give an observer no insight into the probability of that student passing the course. The determining factor of whether a student within a particular group passes the course would be the individual "efforts" that that student put forward to do so.

Roemer points out that looking at measurement of effort alone, even if one could observe such a measurement, would not be enough to distinguish equality of opportunity. Doing so would only take into consideration the direct effects of circumstances. A Roemerian sense of equality of opportunity must take into account indirect effects of circumstance as well. These include changes in individual effort based on the circumstance group that one finds him or

herself in. Consider the previous example, assuming this time that females from two-parent households have a disproportionally high pass rate. Individuals from another circumstance group, males from single-parent households for example, may observe their relatively low probability of passing that particular course and thus adjust their level of effort downward, perhaps reallocating their time and effort to another course in which such a bias might not exist. This reduction or reallocation of effort is not the consequence of choice, per se, but rather the outcome of being in a relatively disadvantaged circumstance group. This indirect effect, according to a Roemerian sense of opportunity equality, should also be compensated for.

The Human Opportunity Index

Based on the Roemerian concept of inequality of opportunity, de Barros et al. (2009) propose the Human Opportunity Index (HOI), a metric that attempts to capture inequality of opportunity within a society. At its core, the HOI measures the differences in access to "basic opportunities" among children. Access among children is used, as access "defines opportunity, because children (unlike adults) cannot be expected to make the efforts needed to access these basic goods by themselves" (p 3). Basic opportunities under this regime are defined as a "subset of goods and services for children, such as access to education, to safe water, or to vaccinations, that are critical in determining opportunity of economic advancement in life. These are either affordable by society at large already, or could be in the near future, given the available technology." (p 3) For their study, the authors chose education and housing opportunities as a baseline measure of opportunities for children. Completion of sixth grade on time, school attendance at ages ten to fourteen, and access to clean water, sanitation, and electricity were selected to measure these baselines. As these are the fundamental factors that compose the HOI, the present study will analyze if these factors, as combined and discounted below, are competent in predicting future levels of income inequality.

The HIO measures the differences in access by combining two measurements of access. The first is the absolute coverage level of a particular opportunity, that is, the mean coverage level. The second measure is a dissimilarity index, or D-index. The D-index separates the population of children into distinct groups based on their life circumstances, such as a child's area of residence or gender. It then assesses the differences in these groups' average access to a particular opportunity to the mean for all groups. Equation 4 in the Methods section of this study offers an computational description of the index. The D-index used in this way "can be interpreted as showing the fraction of all available opportunities that need to be reassigned from better-off groups to worse-off groups to achieve equal opportunity for all" (p 6). Alternatively, it could serve as an indicator as to which groups additional opportunities should be allocated, as they become available, to equalize opportunity. de Barros et al. (2009) combines these two measures, proposing:

$$O = \bar{p}(1 - D), \tag{1}$$

Equation 1: Source de Barros et al. (2009)

where O is the Human Opportunity Index, $\bar{p}$ is the mean access level to opportunities considered within a particular country or region, and D is the dissimilarity index$(0 \leq D \leq 1)$. Equation 1 indicates the case where only one opportunity is considered. Under this proposed metric, the mean availability of opportunities drives the HOI, and the D-index discounts the mean based on the level of inequality in access to opportunities. The HOI will thus be a number between 0 and 1. An HOI of 1 would indicate that there is one hundred percent coverage in all opportunities, and as such, perfectly equal distribution of opportunity. Conversely, an HOI of 0 would indicate

that there is a zero percent coverage rate. In this case, there would also be perfectly equal distribution of opportunities, that is, no child has access to any of the considered opportunities.

A continuation of the work done by de Barros et al. (2009), Molinas et al. (2010) expands on the theoretical and empirical analysis of the Human Opportunity Index. The authors define the circumstances considered in creating the HOI. Table 1 summarizes both the circumstances and the opportunities (or outcomes) that are measured by the HOI and subsequently used in the present study. The relevant methods and findings of the study are reserved for subsequent sections.

Table 1: Summary of HOI Circumstances and Opportunities

Circumstances	**Opportunities (Outcomes)**
Parent's Education	Completed 6th Grade on Time
Family per capita Income	Attends School (ages 10-14)
Number of Siblings	Availability of Running Water in the Home
The Presence of Both Parents	Availability of Sanitation in the Home
Gender of the Child	Use of Electricity in the Home
Gender of the Household Head	
Urban or Rural Location of Residence	

Source: Molinas et al. (2010)

Inequality of Opportunity and Income Inequality: Review of Literature

Brazil is, historically, recognized as one of the most unequal societies in the world. However, it has also made great strides in improving access to education, housing, and other basic necessities over the past several decades. This, in combination with the regularity and quantity of data created by the Pesquisa Nacional por Amostra de Domicílios (PNAD), has made Brazil and attractive case study for inequality of opportunity.

There is a substantial literature investigating intergenerational mobility in Brazil. Lam (1999) and Paes de Barrow and Lam (1993) focus on educational transfer and assesses and parental characteristics in explaining educational inequality. They attribute all differences not attributable to the parental characteristics to individual differences in motivation or effort. Horowitz and Souza (2011) consider intra-household allocation of educational resources within multiple child households. They find declining relationship between intra-household educational attainment and income, indicating that transfers to lower income families, which often have to specialize some children in labor and others in education, may benefit children's educational prospects. Ferreira and Veloso (2006) use data from the 1996 PNAD to estimate the intergeneration mobility of wages between fathers and sons and found that the "degree of intergenerational mobility of wages in Brazil is lower than the one observed in developed countries" (p 182). The authors also found intra-national disparities in intergenerational wage mobility. Sons in wealthier regions (in the Southeast) had higher mobility than their counterparts in poorer regions (in the Northeast). Disparities were also found along racial lines. Dunn (2007) takes a similar approach using 25-34 year old males. He finds an upper-bound elasticity of .69, which he indicates is greater than "any country previously studied" (p 1).

Both Ferreira and Veloso's (2006) and Dunn's (2007) estimates create measurements of intergenerational wage mobility. As Bourguignon, Ferreira and Menéndez's (2007) point out, this measure of intergenerational wage mobility "would only correspond to a measure of inequality of opportunity under the clearly restrictive assumption that parental earnings is a sufficient statistic for all observed circumstances" (p 598). Consider Lam (1999) and Paes de

Barrow and Lam (1993) as described above. They considered several non-wage parental attributes that affected their children's educational mobility. As education is an important factor in earnings outcomes, these non-wage parental factors constitute a circumstance outside of the scope of intergenerational wage mobility, as found by Ferreira and Veloso (2006) and Dunn (2007). The concept of intergenerational wage mobility, thus, cannot be considered a substitute for the concept of equality of opportunity.

Working directly from a Roemerian concept of equality of opportunity, Bourguignon, Ferreira and Menéndez's (2007) use the same 1996 PNAD data as Ferreira and Veloso (2006), selecting five "circumstance" variables: father's and mother's education, race, region of birth, and father's occupation. They estimate the effect that these circumstances, which they describe as "opportunity-forming" (pg 585), have on income inequality in Brazil. They find that these five circumstances explain between ten and thirty-seven percent of the Theil index. Decomposition showed that about sixty percent of the effect on earnings was direct, while the remaining forty percent of the effect was indirect from differentiated efforts. Further decomposition shows that parental education was the most important circumstance within the five selected. While related, Bourguignon, Ferreira and Menéndez's (2007) approach is distinct from the present study in several ways. Primarily, theirs uses one year's data (the 1996 PNAD) and breaks it into seven age cohorts. This study only considers equality of opportunity for children under the age of sixteen. This simplifies the access/opportunity question, as describe previously, that may otherwise be present. Second, this study will use a synthetic measure of inequality of opportunity, which combines and discounts all selected circumstances into one number.

Núñez and Tartakowsky (2011) based a study in Chile on Bourguignon, Ferreira and Menéndez's (2007) work in Brazil. Using data from Chile's 2006 National Socio-Economic Characterization Survey, the authors found that equalizing circumstances to the mean distribution level would produce a reduction of ten to twenty percent in Chile's Gini coefficient. Similar to the findings of Bourguignon, Ferreira and Menéndez (2007), about fifty percent of the change in income inequality was direct, while the other half was indirect thought education accumulation. The authors make the conclusion that "most of the measured inequality is unrelated to heterogeneity in observed circumstances" (p 363). That is, the majority of the difference between incomes is not based on observed "circumstance", but rather, on unobserved "effort". Based on this conclusion, the authors point out that, "Equality of opportunity is thus likely to coexist with a significant amount of observed income inequality, which suggests that promoting equality of outcomes would demand not only equalizing circumstances and opportunities across individuals, but also a dose of pure redistributive policies" (p 363). This highlights an important concept within the context of equity of opportunity discussion: equity of opportunity does not guarantee, or even imply, equity of outcome.

Perhaps the most thorough empirical examination of the concept of equality of opportunity and income inequality to date, however, examines these topics in Sweden, one of the most equal societies in the world. Björklund, Jäntti, and Roemer (2011) exploit a combination of the Swedish Mulit-generational register, two bi-decennial censuses from 1960 and 1980, Statistics Sweden's income register, and the Swedish Military Enlistment Battery to create a set of wide-ranging circumstances, which the authors assert should not affect outcomes, that could nonetheless affect long-term income inequality. The authors use parental income during

childhood, parental education, IQ during adolescence[1], number of siblings, body mass index during adolescence, and family structure. They combined these factors in such a way to create 1152 "types" or homogeneous circumstance groups for Swedish men born between 1955 and 1976.

The authors found that parental income during childhood, an individual's IQ, and the type heterogeneity of effort[2] were the most important factors in determining overall inequality. However, they found that over seventy percent of the income inequality in Sweden could be attributed to "efforts", or more precisely, the error term. This, they assert, is the mark of a highly developed economy and institutions, pointing to de Barros' et al. (2009) findings that thirty to fifty percent of inequality in Latin America is from circumstances. This difference is made even greater when considering the comparative paucity of variables used in de Barros' et al. (2009) study.

Indeed, it is the vast data resources available and employed by the authors that weaken the applicability of this study in other countries. Given the relative lack of viable data sources in most poorer and developing countries, for which inequality of opportunity and income inequality would likely be of most interest, it would be impractical to recreate Björklund, Jäntti, and Roemer's (2011) study on a broader international scale. The present study takes its cue from both of these studies; it uses the HOI metric proposed by de Barros et al. (2009), which is more versatile in the countries for which sufficient and adequate data is available, and only considers the opportunity distributions of children. It also attempts to link the data to future income inequality, as was attempted by Björklund, Jäntti, and Roemer's (2011).

Molinas et al. (2010) is the most recent update of de Barros et al. (2009), assessing the HOI for nineteen countries from in Latin America and the Caribbean for 1995 and 2010 (or the nationally representative survey data closest to those dates). Molinas et al.'s (2010) calculations for Brazil in 1995 and 2008 are summarized in Table 2.

Table 2 Brazilian HOI 1995, 2008
Source Molinas et al. (2010)

	HOI	Coverage Rate	Penalty
1995	57	66	9
2008	76	80	5

The authors' calculations show a significant gain in opportunity expansion as well as in opportunity equalization. Over the thirteen year period, there was an average annual growth of 1.44 HOI points. The present study will examine the specific year-to-year movements over part of this time period.

[1] The authors address possible objections to the inclusion of IQ as a circumstance variable. They argue that even though individual efforts may play a role in one's IQ, that those efforts, prior to a particular age, are not something that an individual can be held responsible or accountable for. Thus, IQ, in the context of an adolescent, can still be considered a circumstance, in the sense that it is not something that should affect future earnings.

2 The authors define heterogeneity of effort as the indirect effect of inequality of opportunity, or the change in effort based on an individual's homogeneous type category.

Data

The data used for this study was taken from the 2001-2008 Pesquisa Nacional por Amostra de Domicílios (National Household survey) or PNAD, collected by the Brazilian Institute of Geography and Statistics, which serves many of the same functions as the United States Census Bureau. The annual survey (except for years in which a census is taken) collects data about the household (the structure itself, location, facilities, etc.), the family structure (number of inhabitants, single or two parent household, total family income, etc.), and individuals within the household (age, race, birthplace, educational history, position within family, etc.). In addition to the numerous biographical and socioeconomic data that are collected every year, each edition of the PNAD contains a supplemental question set that inquires about specialized topics. While some of these supplements would provide additional relevant data for the present study, notably the 2006 PNAD supplement which contained information about the head of household's and spouse's father's occupation, the non-continuous nature of the questions makes their use impractical for the selected methods.

Table 3 indicates the number of households and individuals surveyed in the 2001-2008 editions of the PNAD. From the data about individuals, only those sixteen years old or younger were used to calculate the HOI.

Table 3: Summary of Data Totals

Year	Number of Households	Number of Individuals	Under 16
2008	150,591	391,868	113,171
2007	147,851	399,964	118,478
2006	145,547	410,241	123,955
2005	142,471	408,148	126,101
2004	139,157	399,354	126,055
2003	133,255	384,834	122,236
2002	129,705	385,431	123,424
2001	126,858	378,837	125,405

In order to create the final data set needed to calculate the Human Opportunity Index, the individual level and the household level data were merged via a unique household code present in both data sets. This allowed for household characteristics to be associated with the individuals living in those households.

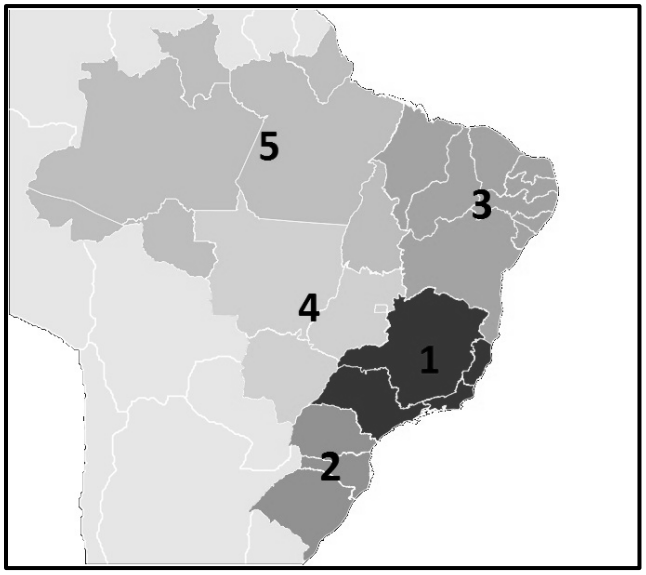
Using the merged data set, variables corresponding to all of the circumstances and opportunities that are included in Table 1 were created either through a direct question from the PNAD, a question deemed an appropriate proxy, or through a household analysis. Direct questions addressed the family per capita income (or a simple calculation for earlier editions of the PNAD), the presence of both parents, the gender of the child, the location of the residence (urban or rural), the completion of sixth grade (only children thirteen to sixteen years of age were considered[3]), school attendance (only children ages ten to fourteen were considered), the availability of running water in the home, and the availability of sanitation in the home. Using

[3] This is in contrast de Barros et al. (2009) and Molinas et al. (2010), which used a regression model to determine the probability of completing the sixth grade on time, and included all children.

the unique household identifier, the number of siblings in a given home was computed. The same method was used to find the gender of the head of household. The education of the head of household was used as a proxy for the parent's education, and the type of lighting used in the home (electric, gas, etc.) was used to determine the presence of electricity.

In addition to these variables, the age of the child was retained (as to only include children in the appropriate age ranges for education opportunities) as well as a variable indicating in which state the household was in (including the Federal Capital). The state variable was subsequently converted into a macroregional variable. The five macroregions designated by the Brazilian Institute of Geography and Statistics, Southeast, South, Northeast, Central West, and North, were used for analysis. Figure 1 and Table 4 indicate the regional variables used.

Figure 1 Numbered Macroregions of Brazil



Source: Menegaz, adapted

Table 4: Macroregional Variables

Number	Region	States
1	Southeast	Espírito Santo
		Minas Gerais
		Rio de Janeiro
		São Paulo
2	South	Paraná
		Santa Catarina
		Rio Grande do Sul
3	Northeast	Alagoas
		Bahia
		Ceará
		Maranhão
		Paraíba
		Pernambuco
		Piauí
		Rio Grande do Norte
		Sergipe
4	Central West	Goiás
		Mato Grosso
		Mato Grosso do Sul
		Distrito Federal
5	North	Acre
		Amapá
		Amazonas
		Pará
		Rondônia
		Roraima
		Tocantins

As with most survey data, missing variables were present across all of the editions of the PNAD used. For this study, individuals who were missing one or more circumstance variables, as defined by Table 1, were excluded from the data set. Those that were missing one or more of the opportunity variables were excluded only from the analysis of the variables for which their data was missing.

The wholesale exclusion of individuals missing these variables likely introduces bias into the model, as non-response is likely to be non-random; however, an inference can be made as to the direction of bias introduced through non-response. There was no non-response for (or a uniform ability to calculate) number of siblings and urban/rural location of residence. It is presumed that non-response would be higher for parental education when education level is low; higher for family structure when both parents are not present; higher when the child is a girl; and higher when the gender of the head of household is female. All of these situations would cause

children with these circumstances not to be considered. Children in these circumstances would almost certainly be less likely, on the whole, to have access to one or more of the opportunities considered. Exclusion would introduce a positive bias to the HOIs calculated.

Non-response is presumably higher for all of the opportunities outlined in Table 1 if a child did not have access to them. Not considering children for which this data is not available would also introduce a positive bias. Overall, there is likely an upward bias of unknown magnitude to the HOIs calculated. The direction and magnitude of the bias, however, is not as important as the consistency of it. As the analysis of the data will be examining trends over time, providing the bias is consistent, bias should not affect the analysis.

In addition, the same positive bias may be among the macroregional data; although, this bias may not be uniform across macroregions. There is likely more bias introduced by non-response in more rural and more impoverished macroregions. Thus, by the same rational as used above, the positive bias may be greater in these areas. This, however, would only be more problematic than the national HOI when making interregional comparisons. Again, providing the bias is consistent, bias should not affect an analysis comparing an individual macroregion to itself over time.

In addition to the datasets created for calculating the HOI, data sets were created to calculate a Gini Index. In order to generate this data set, the total household income for the head of household for each household was extracted from the data set of individuals. Those heads of household that did not report family income were excluded. It is likely that those with a high income would be less likely to report income because of tax concerns or related issues. This would introduce a downward bias in the Gini coefficient. As with the HOI, provided there is consistent bias, the analysis should not be affected.

Methods

Using the data sets described above, the Human Opportunity Index was calculated via similar methods to those set out in Molinas et al. (2010). A logistic model was constructed assessing the probability that an individual child, child *i*, had access to an opportunity, as outlined in Table 1, based on his or her circumstance variables. For completion of sixth grade, only children thirteen and older were included in the regression. For school attendance, only children that were between the ages of ten and fourteen were included. Considering opportunity variables, family per capita income was transformed by the natural log function before included in the regression. All other variables were taken to be categorical.

Using the estimated coefficients, the predicted probability of access to an opportunity was calculated by Equation 2:

$$\widehat{p_i} = \frac{1}{1 + \exp(\widehat{\beta_0} + \sum_{k=1}^{m} \widehat{\beta_k})} \qquad 2$$

Equation 2: Source Molinas et al. (2010)

where $\widehat{p_i}$ is the predicted probability of access to an opportunity for child *i*, $\widehat{\beta_0}$ is the estimated intercept, and $\widehat{\beta_k}$ is the estimated coefficient for circumstance *k*. The overall coverage rate of predicted access was then generated by Equation 3:

$$C = \sum_{i=1}^{n} w_i \hat{p}_i \quad (3)$$

Equation 3: Source Molinas et al. (2010)

where *C* is the overall coverage rate of predicted access, and w_i is equal to $1/n$. Using this coverage rate, the dissimilarity index was calculated via Equation 4:

$$\hat{D} = \frac{1}{2C} \sum_{i=1}^{n} w_i |\hat{p}_i - C| \quad (4)$$

Equation 4: Source Molinas et al. (2010)

where $\hat{D}$ is the estimated dissimilarity index. The penalty for unequal access based on the circumstances considered was then calculated by multiplying *C* with $\hat{D}$.

$$P = C * \hat{D} \quad (5)$$

Equation 5: Source Molinas et al. (2010)

The HOI was then computed by subtracting the penalty from the overall coverage rate of predicted access.

$$HOI = C - P \quad (6)$$

Equation 6: Source Molinas et al. (2010)

Equation 6 is equivalent to Equation 1, proposed by de Barros et al. (2009).

The Human Opportunity Index was computed for each year from 2001-2008 and for each macroregion for 2001, 2005, and 2008. The HOIs are reported in Table 5.

Table 5: National and Regional HOIs 2001-2008

Year	HOI	Region 1	Region 2	Region 3	Region 4	Region 5
2001	70.2	70.4	72.9	69.6	77.3	73.4
2002	72					
2003	71.5					
2004	66.8					
2005	66.5	64.8	67.2	66.9	71.0	68.3
2006	66.7					
2007	54.9					
2008	67	65.9	67.3	66.4	66.3	68.3

A Gini Index, or Gini, was also calculated nationally for each year in the data set and macroregional Ginis for the same years that macroregional HOIs were calculated: 2001, 2005, and 2008. The Gini is a measure of inequality of individuals, household, or other discrete groups.

In this case, it was used to measure the inequality of household incomes within Brazil. The Gini index is a real number between zero and one hundred, with a Gini of one hundred representing a perfectly unequal society (that is, one in which all resources considered belong to one person) and a Gini of zero representing a society in which all of the resources considered are distributed perfectly equally.

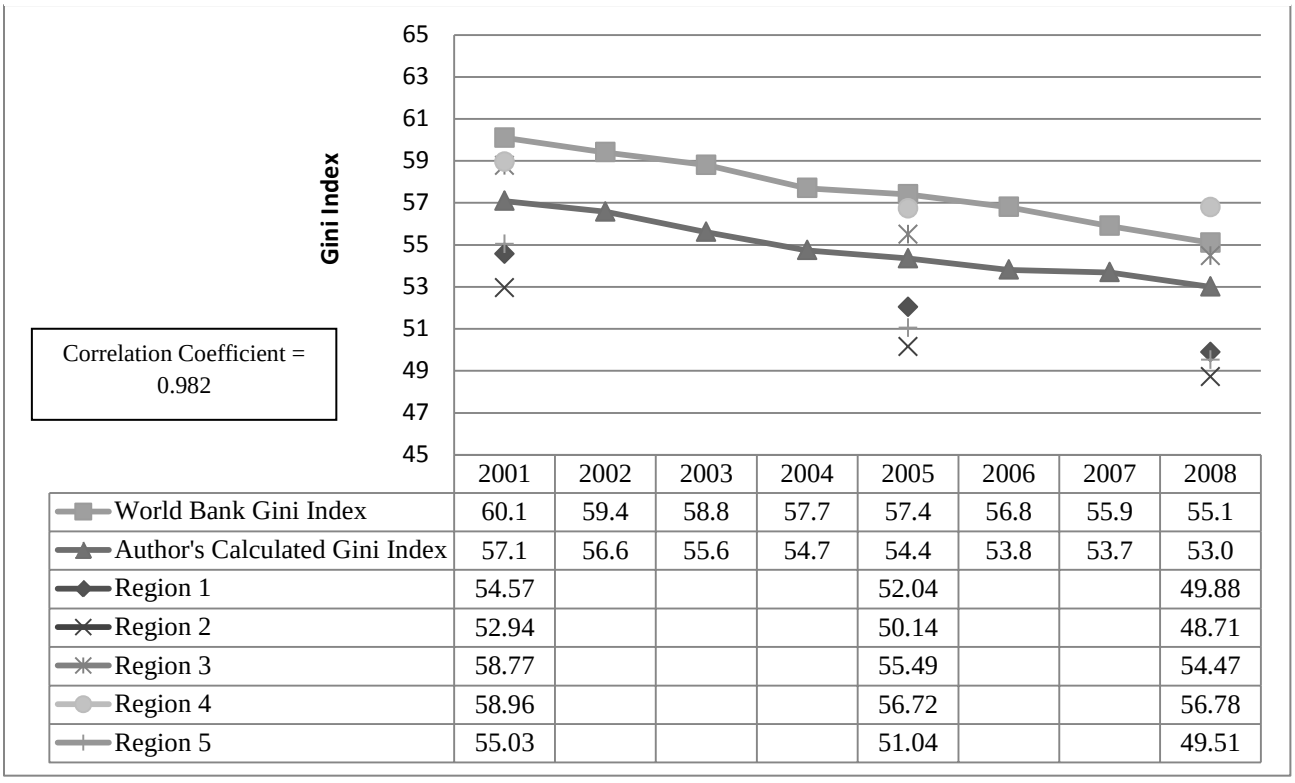
Equation 7, developed by development economist Angus Deaton, was used to calculate the Gini for each year and region using the household income data described in the previous section.

$$G = \frac{N+1}{N-1} - \frac{2}{N(N-1)u}\left(\sum_{i=1}^{n} P_i X_i\right) \qquad 7$$

Equation 7: Angus Deaton (1997, 139)

The results are presented in Figure 2. The data used introduced bias (presumed to be downward), as discussed in the previous section. Indeed, when calculated, the national, annual Gini coefficients calculated from this data is consistently lower (showing less income inequality) than the Gini coefficients for Brazil published by the World Bank. They do, however, have a very high linear correlation coefficient (.982), as shown by Figure 2.

Figure 2: National and Regional Gini Indexes



	2001	2002	2003	2004	2005	2006	2007	2008
World Bank Gini Index	60.1	59.4	58.8	57.7	57.4	56.8	55.9	55.1
Author's Calculated Gini Index	57.1	56.6	55.6	54.7	54.4	53.8	53.7	53.0
Region 1	54.57				52.04			49.88
Region 2	52.94				50.14			48.71
Region 3	58.77				55.49			54.47
Region 4	58.96				56.72			56.78
Region 5	55.03				51.04			49.51

Results

Results and analysis in this section will be disused in three ways: analysis of the HOI and related decompositions, analysis of the Gini Index, and analysis of the relationships between the two indexes.

The HOI was calculated for 2001-2008 nationally, and by macroregion for 2001, 2005, and 2008. The full results are listed and depicted in Figure 3. In light of the expected direction of the bias discussed in the previous section, it is puzzling that the HOI calculated for the present study is lower than that calculated by Molinas et al. (2010) for 2008. However, when the penalty is calculated as a percentage of the coverage rate for both the preset HOI score (5.22%) and Molinas et al.'s (2010) score (6.25%), they are within two percentage points of each other. And, as stated previously, assuming the bias remains constant over time, it should not affect analysis of trends over time.

There appears to be a fairly stable horizontal trend over the eight-year-period studied. There is an abnormal 13.5 point drop from the 2006 HOI to the 2007 HOI, the majority of which is recovered in the 2008 calculation. A simple OLS regression, including only year as the regressor and HOI as the regressand, shows a slightly negative slope, as reported in Equation 8.

$$\widehat{HOI} = 73.74 - 1.21(Year) \quad (8)$$
$$(3.233)\ \ (.6404)$$

Equation 8: Source Author's Calculation

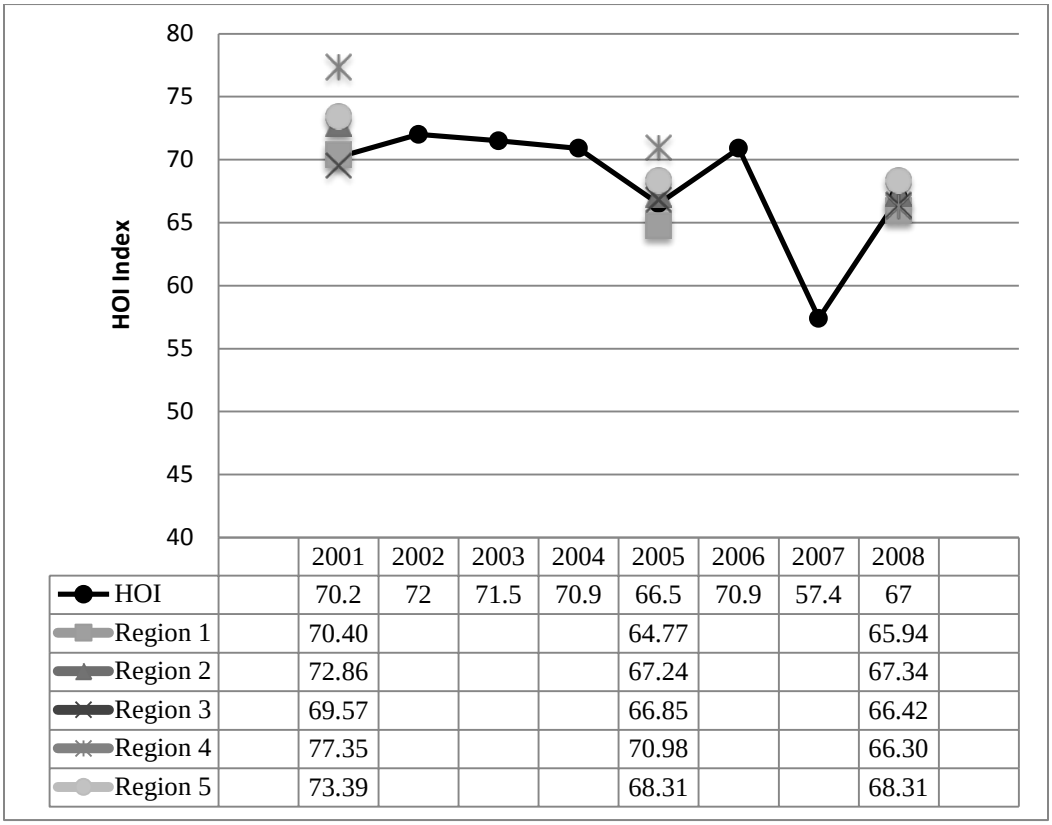
If the abnormally low 2007 data point is removed, the regression is as shown in Equation 9.

$$\widehat{HOI} = 72.19 - .056(Year) \quad (9)$$
$$(1.503)\ \ (.3207)$$

Equation 9: Source Author's Calculation

Both estimated coefficients show a slight downward trend with respect to time. Neither, however, is statistically significant, even at the .10 alpha level. Over such a short time period (when considering substantial societal shifts), it is not prudent to conclude that there is downward movement in the level of inequality of opportunity in Brazil, as measured by this index over the eight years considered.

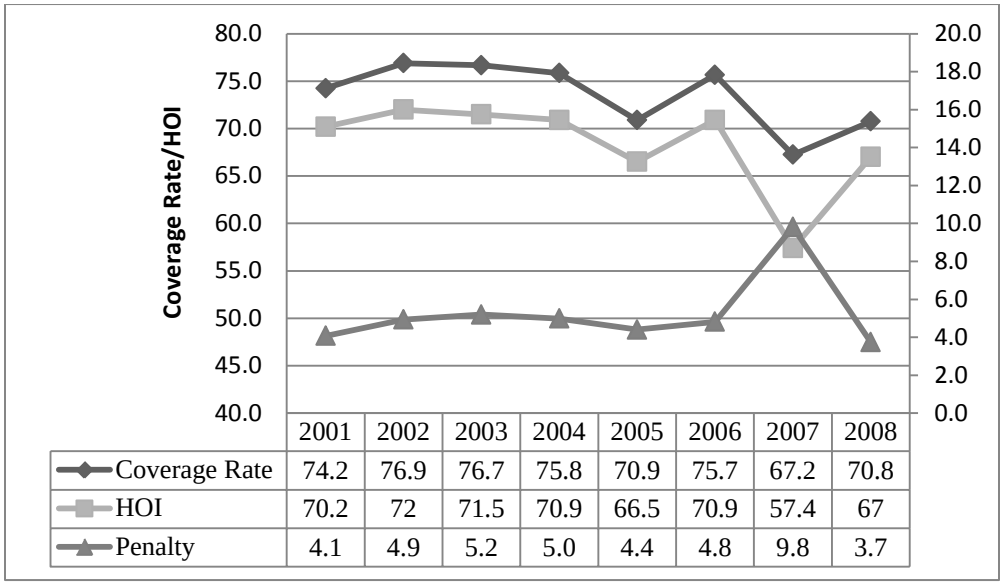
Figure 3 Brazilian HOI 2001-2008



	2001	2002	2003	2004	2005	2006	2007	2008
HOI	70.2	72	71.5	70.9	66.5	70.9	57.4	67
Region 1	70.40				64.77			65.94
Region 2	72.86				67.24			67.34
Region 3	69.57				66.85			66.42
Region 4	77.35				70.98			66.30
Region 5	73.39				68.31			68.31

A decomposition of the HOI, breaking the HOI into the coverage rate of predicted access and the penalty, as shown in Equation 6, is reported in Figure 4. This decomposition indicates that the unusually low score for 2007 was due in larger part to an inequity penalty (9.8) that was 4.6 points higher than the second highest penalty, rather than an unusually low coverage rate of predicted access. The 2007 penalty as a noticeable exception, there is almost no change in the HOI penalty component over the eight year period.

Figure 4 Brazilian HOI Decomposition 2001-2008



	2001	2002	2003	2004	2005	2006	2007	2008
Coverage Rate	74.2	76.9	76.7	75.8	70.9	75.7	67.2	70.8
HOI	70.2	72	71.5	70.9	66.5	70.9	57.4	67
Penalty	4.1	4.9	5.2	5.0	4.4	4.8	9.8	3.7

A more interesting trend can be found in the regional data in 2001, 2005, and 2008. Results from a one-tailed F-test for two sample variances are reported in table five. They show a statistically significant decrease in the variance of HOIs among macroregions between 2001 and 2008 at the .05 alpha level. Almost exactly half of the reduction (49.46%) in variance was produced in the period between 2001 and 2005 and half (50.51%) between 2005 and 2008. This indicates that while there was not a statistically significant trend for the overall HOI during this time period, the opportunity gaps between the macroregions was decreased by a statistically significant amount.

Table 6 F-test for Two Sample HOI Variances

	2001	***2008***
Mean	72.712	66.862
Variance	9.299	0.917
F	10.145	
P(F≤f) one-tail	0.022	
F Critical one-tail	6.388	

Considering the calculated Gini Indexes, there is a monotonic downward trend in both the indexes calculated by the World Bank and for the present research. A simple OLS regression, including only year as the regressor and Gini inex calculated for this research as the regressand, shows a slightly slope, as reported in Equation 10.

$$\widehat{Gini} = 57.47 - .581(Year) \quad 10$$
$$(.219)\ \ (.043)$$

Equation 10: Source Author's Calculation

The coefficient corresponding to the year is significant at a .01 alpha level, and the regression has an R^2 value of .968, showing a strong, negative linear relationship.

Unlike the HOIs, there is no significant shift in the variance among the macroregions. Table 6 reports results from a one-tailed F-test for two sample variances of the macroregional Gini indexes from 2001 and 2008. The data indicates that there was an increase among macroregional income disparities between 2001 and 2008, although not to a statistically significant level.

Table 7 F-test for Two Sample Gini Variances

	2001	***2008***
Mean	56.052	51.870
Variance	7.188	12.606
F	0.570	
P(F≤f) one-tail	0.300	
F Critical one-tail	0.157	

Because of the eight year time frame, results about long term connections between the HOI and income inequality as measured by the Gini index are limited. As the HOI only considers access to educational and housing opportunities for children under seventeen years old, the full effects of changes in opportunity equality would likely not be observed in measurements of income equality would for several decades. Given these caveats, the results from simple OLS regressions, including only the 2001 HOI national and macroregional scores as the regressor and the national and macroregional Gini indexes from 2005 and 2008 as the regressands are reported in Equations 11 and 12.

$$\widehat{Gini\ 2005} = 40.927 + .171(HOI\ 2001) \quad (31.902)\ (.441) \qquad 11$$

Equation 11: Source Author's Calculation

$$\widehat{Gini\ 2008} = 26.890 + .348(HOI\ 2001) \quad (37.794)\ (.522) \qquad 12$$

Equation 12: Source Author's Calculation

Neither coefficient is significant even at an alpha level of .10. Indeed, the sign of the intercept is not consistent with the intuition that broadening opportunity equality among children would decrease future income inequality. Using a linear correlation coefficient, that is reported in Table 8, it is noted that this puzzling, positive correlation is increasing both over the time interval from 2001 to 2005 and from 2005 to 2008.

Table 8 Correlation Coefficient Matrix

	HOI:2001
HOI:2001	1
Gini:2001	0.142
Gini:2005	0.190
Gini:2008	0.316

Discussion

An estimate of the correlation between the Gini Coefficient and the Human Opportunity Index, the primary question of interest, was not able to be estimated. The time period considered, 2001 to 2008, was not sufficiently long to provide enough data to have robust results regarding any associations that connect the two measures. Although not statistically significant, a perplexing pattern did emerge of a direct linear correlation over both the time intervals from 2001 to 2005 and from 2005 to 2008. If this pattern were to continue and become statistically significant at an appropriate alpha level, it would indicate that more opportunities in education and housing were present in Brazil, and as those opportunities were more equitably distributed, the overall income distribution in Brazil would become less equal. The paucity of annual data points produced in this study and relatively short time span considered make it very possible for the relationships found to be incorrect. If, however, future studies were to confirm this direct relationship, it could call into question the usefulness of either of the components of the Human Opportunity Index, the calculation method of the HOI, or both in creating positive impacts in real levels of equity. While it is unlikely that more equal educational access would have a positive impact on future income inequality, the impact of housing may be negligible and clouding the effects of education.

What is perhaps the most significant finding is the decrease in the variance of HOI levels in Brazil over the time period considered. There was a uniform decrease in variance among macroregions over the 2001 to 2005 period and from the 2005 to 2008 period. The total decrease in variance was statistically significant at a .05 alpha level. This decrease in macroregional disparities in opportunity inequality as measured by the HOI is significant in that it addresses a long standing disparity in Brazilian society. Region 1, the southeast region, has historically been the wealthiest and most prosperous region. It is home to the two largest cities in the county, São Paulo and Rio de Janeiro, and was the seat of government when the Portuguese monarchy fled Europe to Brazil. In contrast, Regions 4 and 5, the Central West and North, have historically been largely agriculturally based or completely underdeveloped. It is beyond the scope of the present study to speculate as the cause of this decrease in macroregional disparity; it is, however, encouraging to see a decrease in macroregional disparities.

The HOI only includes an urban/rural designation as a circumstance variable; it does not include a state or regional variable. If it were, however, included in a modified HOI, it would likely have a significant impact for the majority of Brazilian history. However, if the trend of decreasing macroregional disparities in opportunity continues, that likely significance would evaporate.

The primary track for future research regarding the primary question of interest for this study is to calculate both national and macroregional HOI and Gini coefficients for a larger

series of time, ideally two to three decades. Using a longer time series, one could more accurately assess relationships between changing HOIs and changing Gini indexes. Also, this study made the naïve assumption that any relationship between the two indexes would be linear. This is not necessarily, indeed likely not, the case. One could test more appropriate nonlinear models to acutely assess the nature and lag structure of any relationships.

Additional data would also facility an analysis of various HOI decompositions. One could assess national and macroregional trends in coverage levels and dissimilarity penalties as well as HOI decompositions to determine what combination of housing and educational opportunities are responsible of gains or losses in HOI score. These additional insights could be used by policy makers at many levels of government to create policy that could effectively tackle inequality issues in Brazil.

References

Aghion, Philippe, and Patrick Bolton. 1997. A theory of trickle-down growth and development. *Review of Economic Studies* 64 (2) (04): 151-72.

Banerjee, Abhijit V., and Andrew F. Newman. 1993. Occupational choice and the process of development. *Journal of Political Economy* 101 (2) (04): 274-98.

Bourguignon, Francois, Francisco H. G. Ferreira, and Marta Menendez. 2007. Inequality of opportunity in brazil. *Review of Income and Wealth* 53 (4) (12): 585-618.

Björklund, Anders, Markus Jäntti, and John E. Roemer. "Equality of Opportunity and the Distribution of Long-Run Income in Sweden." *IZA Discussion Paper Series*. . http://ideas.repec.org/p/iza/izadps/dp5466.html (accessed April 23, 2012).

de Barros, Ricardo Paes, Francisco H. G. Ferreira, Jose R. Molinas Vega, and Jaime Saavedra Chanduvi eds. 2009. *Measuring inequality of opportunities in latin america and the caribbean*With Mirela de Carvalho, Samuel Franco, Samuel Freije-Rodriguez, and Jeremie Gignoux; Washington, D.C.:; World Bank;; Houndmills, U.K. and New York:; Palgrave Macmillan.

de Barros, Ricardo Paes, and David Lam. 1993. Desigualdade de renda, desigualdade em educacao e escolaridade das criancas no brasil. (with english summary.). *Pesquisa e Planejamento Economico* 23 (2) (08): 191-218.

Deaton, Angus. 1997. *The analysis of household surveys: A microeconometric approach to development policy*Baltimore and London:; Johns Hopkins University Press for the World Bank.

Dunn, Christopher E. 2007. The intergenerational transmission of lifetime earnings: Evidence from brazil. *B.E.Journal of Economic Analysis and Policy: Contributions to Economic Analysis and Policy* 7 (2).

Falk, Armin, Ernst Fehr, and Urs Fischbacher. 2003. On the nature of fair behavior. *Economic Inquiry* 41 (1) (01): 20-6.

Ferreira, Sergio Guimaraes, and Fernando A. Veloso. 2006. Intergenerational mobility of wages in brazil. *Brazilian Review of Econometrics* 26 (2) (11): 181-211.

Galor, Oded, and Joseph Zeira. 1993. Income distribution and macroeconomics. *Review of Economic Studies* 60 (1) (01): 35-52.

Goldstein, Markus, and Christopher Udry. 2008. The profits of power: Land rights and agricultural investment in ghana. *Journal of Political Economy* 116 (6) (December): pp. 981-1022.

Hoffman, Elizabeth, Kevin A. McCabe, and Vernon L. Smith. 1996. On expectations and the monetary stakes in ultimatum games. *International Journal of Game Theory* 25 (3): 289-301.

Horowitz, Andrew W., and Andre Portela Souza. 2011. The impact of parental income on the intra-household distribution of school attainment. *Quarterly Review of Economics and Finance* 51 (1) (02): 1-18.

Lam, David, and Suzanne Duryea. 1999. Effects of schooling on fertility, labor supply and investments in children, with evidence from brazil. *Journal of Human Resources* 34 (1) (Winter): 160-92.

Menegaz, Felipe. Wikimedia Commons, Last modified June 11, 2007. Accessed April 23, 2012. http://upload.wikimedia.org/wikipedia/commons/f/f6/Brazil_Labelled_Map.svg. Adapted.

Molinas Vega, Jose R., Ricardo Paes de Barros, Jaime Saavedra, and Marcelo Giugale. 2010. *Do our children have chance?:The 2010 human opportunity report for latin america and the caribbean.* Washington, D.C.: The International Bank for Reconstruction and Development/The World Bank, .

Nunez, Javier, and Andrea Tartakowsky. 2011. The relationship between income inequality and inequality of opportunities in a high-inequality country: The case of chile. *Applied Economics Letters* 18 (4-6) (March): 359-69.

Rawls, J. (1999). *A theory of justice*. Oxford: Oxford University Press.

Roemer, J. E. (1998). *Equality of opportunity*. Cambridge, Mass: Harvard University Press.

Sen, A. K. (1999). *Commodities and capabilities*. Delhi: Oxford University Press.

Michael Walton and Francisco H. G. Ferreira, "Inequality of Opportunity and Economic Development" The World Bank, Policy Research Working Paper Series: 3816, 2006), http://0-search.ebscohost.com.library.uark.edu/login.aspx?direct=true&db=ecn&AN=0814360&site=ehost-live; http://www-wds.worldbank.org/servlet/WDSContentServer/WDSP/IB/2006/01/12/000016406_20060112091459/Rendered/PDF/wps3816.pdf.