MEASURING INEQUALITY OF OPPORTUNITIES FOR CHILDREN

Ricardo Paes de Barros, Jose R. Molinas Vega, and Jaime Saavedra
(with collaboration of Mirela Carvallo, and Samuel Franco and Samuel Freije).

1. INEQUALITY OF OPPORTUNITY: CONCEPTS AND ANALYTICAL FRAMEWORK

1.1. Concepts

Types of inequality

Even though absolute poverty and inequality are closely related concepts, the aims to reduce poverty and inequality have received clearly different supports. While the call for reductions in absolute poverty has always received universal and unconditional support, being the first of the Millennium Development Goals, the call for reduction in inequality has never received unconditional support.

The support for inequality reduction seems to vary among groups, as well as according to the level and nature of inequality. When inequality is extremely high, as in most Latin American countries, and closely related to family origin, gender and ethnicity, consensus that inequality must be reduced can, in general, be more easily achieved. However, when inequality is the result of differences in effort, choices or even talents among persons who had access to the same opportunities and received the same treatment, then the support for compensatory policies to reduce inequality is usually much weaker.

Thus, in the debate about public policy and inequality reduction we must recognize that inequality is made of heterogeneous components, some much more unjust, undesirable and unnecessary than others. Even though we could easily reach consensus toward policies devoted to reduce or eliminate some of these components, consensus for fighting other sources might never be achieved. Some potential sources of inequality may be necessary to give everyone proper incentives (inequality caused by differential efforts). Some may be tolerated (like the inequality due to differences in talents), in particular, when attempts to reduce it could interfere with other ethical objectives, like privacy.

Hence, for the debate about the relation between public policy and inequality reduction to advance, it is necessary to decompose overall inequality and evaluate the magnitude of its components. Diagram 1 presents a basic framework to decompose inequality by source and to relate the effort to reduce it to public policy. Next, we use this diagram to elaborate on this breakdown of the overall inequality level.

---

1 Diagram 1 was constructed with the objective of providing a synthesis of inequality decomposition, presented in this section, and its relation to public policy.
Choices and circumstances

To a larger extent, there are two types of determinants on a person’s welfare. On one hand, we have those factors that could have been different if the person had chosen different paths, which were actually available to her. In this case, individuals with identical choice sets reach different outcomes only due to differences in their choices. Hence, one may argue that, in these situations, they must, themselves, be held accountable or responsible for the resulting inequality.

On the other hand, we have differences in welfare among individuals who would have made the same choices if they had had the same opportunities. In this case, all the differences are due to exogenous factors that lead individuals to face different opportunity sets. Contrary to the previous situation, now one must argue that they cannot possibly be accountable or responsible for the resulting inequality. We usually refer to this component as the inequality of opportunity.

Although many would argue that this source of inequality is intrinsically unjust, not everyone would unconditionally agree that public policy must be used to reduce it. The inequality of opportunity, as well as the overall inequality, is made of heterogeneous components that must be disentangled before we can reach a consensus on its justifiability, desirability and on how much we think public policy should be used to reduced it.

Outcomes and Circumstances
The very concepts of outcomes and circumstances depend on context. If we define inequality of opportunity as the inequality of outcomes due to factors beyond individual control, then it would be natural to define circumstances as the factors (beyond individual control) and outcomes as advantages that are influenced by them, but we wish they were not.

In principle, outcomes themselves and their relation to circumstances, in particular, do not need to be influenced by social choice. However, if the measurement of inequality of opportunity is to have any policy relevance, outcome level and its relation to circumstances must at least conceptually be influenced by social choice.

Hence, we define circumstances as factors beyond individual control (like race) that should not but that actually do affect outcomes of interest (like wage). Outcomes of interest are advantages (like wages) that can be modified by social choice (like subsidized education). Moreover, outcomes (like wage) are in general dependent on circumstances (like race) with this dependency, however, being, at least conceptually, modifiable by social choice (like affirmative action programs). It is important to emphasize that outcomes should depend on social choice through two channels. Social choice must conceptually be able to modify both the level of the outcome and its dependence on circumstances, see Figure 1.

**Figure 1: The relationship between Outcomes, Circumstances and the role of Social Choice**
Second chance and preference for equity

The fact that inequality caused by choice may be justifiable does not necessarily prevent it from being undesirable. Even though we should, due to other competing principles, agree that everyone has the right to make their own choices, nothing prevent us to cheer for better choices. The freedom of choice does not prevent the society for preferring that everyone makes the choices that are best for them. Hence, the resulting inequality in welfare, although justifiable, may not be considered desirable or even necessary.

Support for public policies devoted to give a second chance to those who recognize having not chosen properly the first time are evidence that, although justifiable, society is still interested in devoting resources to reduced inequality caused by differences in choices. Society does not seem indifferent to inequality, even when it is entirely caused by differences in choices between equally well informed individuals with identical opportunity sets.

Types of inequality of opportunity

Even when we constrain ourselves to exogenous sources of inequality, not all of them are perceived as being equally undesirable. Hence, to better design public policy it is also fundamental to isolate and measure each one of the components of the inequality of opportunity. In this case it may be important to differentiate three types of factors: intrinsic and personal characteristics, discriminatory treatment and the access to social services.

First, we have the factors that, although still exogenous, are intrinsic to the person. These are mainly biological factors (genetic inheritance and genetic lottery) and post-born daily luck. To the extent that these factors lead to differences in productivity, achievement, results or virtue (i.e., some being preferred to others as a function of what they can perform or provide), in a meritocratic environment in which we give to everyone what they deserve, the inequality so generated may be considered, if not just, at least acceptable.

A recurrent problem with this source of inequality is how much advantage or recognition should be assigned to productivity, achievement or virtue. A meritocratic society may be perceived by some as valuing same talents too much, or it may be criticized by valuing just a very narrow set of talents. Above all, for the inequality among unequally talented individual to be considered fair, we need first to demonstrate that the values we are attributing to talents are just.

Ideally, we would prefer to live in a society in which, despite all human diversity, every person is equally deserving. In other words, we would prefer societies where a wide variety of talents and virtues are valued. In this case, even though individuals may be very different regarding their talents, talents’ values are such that, at the end, everyone is equally productive and deserving. Even those strongly in favor of a meritocratic
environment would certainly dislike a society where only one type of virtue has value and only a small fraction of the population has this virtue.

Although usually public policy has not been extensively used to modify the value of talents and virtues, it can, and probably should, be used with this role. It may be an important role for public policy to broaden the use and value of all talents (integrated and inclusive schools may be an important policy in this group). It is certainly extremely important to have a public policy that celebrates human diversity and values all types of talents. Broadening the set of talents with value might lead to important reductions in inequality of welfare.

In the next two topics we elaborate on the other two components of the inequality of opportunity

*Inequality of treatment and discrimination*

But the inequality of opportunity does not result only from the inequality due to deserving traits in a pure meritocratic environment. Quite often, equally talented and productive individuals are treated differently, receiving different access to the best jobs or receiving different wages even when performing the same tasks. In this case, inequality is generated by the unequal treatment of equally deserving individuals. It constitutes a violation of both meritocratic principles and equal opportunity values. We usually refer to this unequal treatment of equals as discrimination.

Consensus on the unfairness of this source of inequality and on the fact that it is avoidable and unnecessary seems easy to be achieved. Nevertheless, the amount of resources societies are willing to allocate to eliminate this source of inequality could still vary considerably and be open to debate.

There are at least two worth mentioning channels through which discrimination could be generated. In the first one, it operates through the unequal access to social services (like education or health). Due to this unequal access, inequality in valuable acquired characteristics, such as formal education, might be produced even among equally talented persons. This inequality in valuable traits will lead to unequal outcomes even in societies where everyone receives what they deserve (i.e., outcome-meritocratic societies). Naturally, the problem with these societies is the lack of equal opportunity to acquire the valuable characteristics that are being so meritocratically translated in higher desirable outcomes (income and welfare, for instance).

Through the second channel, discrimination operates after the access to services is granted. In this case, even when everyone is treated equally when accessing relevant social services, differential treatment is received during the service, leading to differential benefits. Different ethnic groups may be equally treated for admission to a school, but discriminated while studying. Likewise, different genders may be equally treated when considered for a job, but they may be paid differently for performing the same task.
People dressing differently might get equal access to a restaurant, but they might not receive the same treatment by the waiter.

Through either of the two channels formal equality of opportunity is violated, since equally talented and motivated individuals are being treated differently, leading to differential access to opportunities and to better positions. Anti-discriminatory policies would be the appropriate and necessary measure in these situations.

*Inequality of access to social services*

A third component of the inequality of opportunity results from the unequal access to social services, due to differences in family resources. In this case, children from poor parents are not being discriminated, they just do not have equal access to services, considered important for them, to develop and fully utilize their talents, just because their families lack the necessary resources. In fact, one may be hungry just because one does not have the necessary resources to buy food.

Lack of resources may impair not just the access to social services but also the ability to benefit from these services. Children from poor parents may have both: more limited access to schools and learning disadvantages. They may not have books at home or they may have parents with greater difficulty to read for them.

Whenever the access or the ability to benefit from social services depends on family resources, the ideal of equal opportunity is violated and social immobility is generated. In this case, equally talented children from different social backgrounds are not going to have the same opportunities and outcomes for reasons outside their control.

The role of public policy in this case is unanimously recognized. Equal access to social services must be provided to all. Equal conditions for equally benefiting from these services must also be guaranteed to all. The natural implications are (a) subsidized access to social services for all, or at least for the poor, and (b) a guaranteed minimum income to ensure that all have the necessary conditions for benefiting from available services.

*The central role of the inequality of access to social services and the equal opportunity to benefit from them*

As we have emphasized in the previous section, for getting support and for designing adequate public policies aimed at reducing inequality it is important to disentangle the overall inequality into its components. Among them, the unequal access to social services and the unequal opportunity to benefit from these services, due to discrimination or differences in family resources, are perhaps the components whose reductions command the greatest consensus. They are also, probably, the components more suitable for direct public policy interventions.

To the extent unequal access to social services and unequal opportunity to benefit from them are due to exogenous factors (like gender, ethnicity, place of origin or family
background), they are not just part of the overall inequality, but certainly also a constituent part of the inequality of opportunity. Even though the concept and measurement of inequality of opportunity may remain very debatable, there are no doubts that all the unequal access to social services and the unequal opportunity to benefit from them, due to differences in family resources and discrimination, are integral parts of it. This is particularly true when we concentrate the attention on children. For them, family resources and composition are certainly exogenous.

Moreover, even among the components of the inequality of opportunity, the unequal access to social services and the unequal opportunity to benefit from them, due to discrimination and lack of family resources, play a major and central role. In fact, although the complete endogeneity of choice (are people really free to choose? All of them are equally informed and have the same opportunity set?) and the fairness of rewarding inherited talents are always under questioning, nobody would question the unfairness of children’s unequal access and opportunity to fully benefit from social services due to discrimination or differences in family resources.

1.2. FROM CONCEPTS TO MEASUREMENT

Inequality of opportunity has always received great attention in the policy debate as well as by philosophers. Nevertheless, empirically it does not seem to have a universal definition. It has traditionally been measured as the association between family background and children’s outcomes by sociologists. Recently, economists have been given growing attention to the measurement of inequality of opportunity for continuous outcomes like income. In this section we seek to review the main notions of inequality of opportunity important for empirical and applied studies.

Defining an empirically tractable notion of inequality of opportunity

Let \( y \) be an individual outcome of general interest: education, income or access to an important social service, for instance. Let us divide its determinants in two broad groups. One group made of all factors beyond individual control. Let \( c \) denote a vector with all these determinants. We refer to them as circumstances. The other group is formed by all factors under individual control. Let \( e \) be a vector with all these determinants. We refer to them as choices or efforts. Individuals are considered fully responsible for choices and effort.

---

2 For recent reviews of this literature in economics, see Paragine (1999) and Fleurbaey and Maniquet (2004). For classical treatments in the philosophical literature, see Dworkin (1981a,b), Cohen (1989) and Arneson (1990). For excellent recent discussions on the concept of equality of opportunity see Hild and Voorhoeve (2004) and Fleurbaey (2005).

3 See, for instance, the classical work of Boudon (1974).

Even though this dichotomist decomposition is widely used, it is not without major
difficulties. Surely, factors outside individual control are easy to define and to recognize,
choices however are much more difficult to classify. Suppose the choice set for two
persons are completely distinct. Their choices, as a consequence, ought to be quite
different too. Would we consider the differences in the choices they made a result of
personal decisions or a result of exogenous differences in the choice sets? If, due to
factors beyond individual control, the choice sets are quite different, then it will be
awkward to consider the individuals fully responsible for the choices they made.

Let \( \psi \) be the function relating the outcome to its determinants. Hence,
\( y = \psi(c, e) \) and the outcome distribution and its corresponding degree of inequality
are a function of \( \psi \) and the joint distribution of its determinants \((c, e)\).

**Equality of opportunity as the stochastic independence of outcome and circumstances**

Under perfect equality of opportunity, circumstances should not have any impact on
outcome. More precisely, the outcome, \( y \), and the circumstances, \( c \), would ideally be
stochastically independent, i.e., the distribution of \( y \) conditional on \( c \) should be the same
for all possible values of \( c \). If \( y \) were income and \( c \), race, equal opportunity would require
the distribution of income among one race to be identical to the distribution among other
races. Under this definition, a measure of inequality of opportunity would be a measure
of the differences between the outcome conditional distributions. If race were the unique
circumstance, inequality in opportunity would be a measure of the distance from the
outcome distribution for blacks to the corresponding distribution for whites, for instance.

It should be clearly understood, however, that the distance between these conditional
distributions and, hence the magnitude of the inequality of opportunity, is determined by
the shape of the function \( \psi \) as well as by the stochastic dependence between effort and
circumstances. In fact, for the ideal of equal opportunity to be achieved, it is necessary (a)
that circumstances are not really among the outcome determinants and (b) that effort and
circumstances are not stochastically dependent. For instance, if income is a function of
race and effort, then the distance between the distribution of income among white and
blacks will depend on race, in part, to the extent that race has a direct impact on income
and in part to the extent that the distribution of effort is different in the two groups.

To see why the stochastic independence of effort and circumstances is not sufficient,
let \( y = c + u \). In this case, \( F_y(t|c) = F_y(t-c) \). The outcome conditional distribution
depends on \( c \). To see why \( \psi \) being invariant with \( c \) is not sufficient, let \( y = u \). In this case,
\( F_y(t|c) = F_y(t|c) \) and the outcome conditional distribution depends on \( c \), as long as effort
and circumstances are not stochastically independent.

**Equality of opportunity as equal outcome for equal effort**
There is, however, an alternative definition of equal opportunity were the stochastic dependence of effort on circumstances is not included. In this alternative we reverse the conditioning and consider the distribution of outcomes conditional on effort. Equality of opportunity would imply perfect equality of outcomes given effort, i.e., equal outcome for equal effort. In this case the inequality of opportunity should be measured not as a contrast between outcome distributions conditional on circumstances, as in the previous case, but directly as the inequality in the outcome distributions conditional on effort. Naturally, there will be many measures of inequality in opportunity, one for each choice of the level of effort.

The two approaches are not equivalent at all, and the choice between them poses serious ethical questions. After all, the dependence of effort on circumstances is a departure from the equal opportunity ideal. Only when effort is stochastically independent of circumstances the two approaches would be identical.

Before going into measurement issues it should be highlighted that effort being a choice does not imply that all differences in outcome due to effort are fair or compatible with the equality of opportunity. For instance, people may be discriminated based on their choices and nobody would consider this situation as fair. If due to pure discrimination, the wage of a worker is lower as a result of his choice of religion, sexual orientation, local of residence (favela, for instance) or the race of his/her spouse, then we have a situation where an important outcome depends on choices in a clearly unfair way. Hence, for all outcome differences due to efforts to be compatible with equality of opportunity, we have to be sure that the relation between effort and outcome is perfectly meritocratic.

Clearly, these two concepts of equal opportunity have quite different informational requirements to be measured. In the first approach, measures could be obtained without effort being observed and the function $\psi$ being known or estimated. In the second approach, however, either effort must be observed or the function $\psi$ needs to be estimated. But, how $\psi$ could be estimated when effort is unobserved and stochastically dependent on circumstances?

### 1.3. UNOBSERVED CIRCUMSTANCES AND STOCHASTICALLY DEPENDENT EFFORT

In order to construct a practical index of inequality of opportunity we are going to proceed as if (a) the relation between log-outcome, circumstance and effort were additively separable, (b) effort were stochastically independent of circumstances, (c) the relation between effort and outcome were meritocratic, and (d) all circumstances were observed. Surely, none of these conditions are strictly valid in the real world.

Unfortunately, we do not have information necessary to adjust the index to these departures. All we can actually do is to clarify the consequences of these departures and stress the care that one should have when using this index for policy recommendations. The next two sections are dedicated to identify the problems and highlight their potential consequences.
As before, let $y$ be an observed outcome and $c$ a vector with all circumstances. Since we do not observed all of the circumstances and some are more prone to policy intervention than others, we are going to concentrate on a set of observed socially-determined circumstances. Let $x$ denote a vector with all these observed socially-determined circumstances. We are going to treat as socially determined all the circumstances that are not intrinsic to the person. They result from differences on how a person is socially treated. Examples are differences related to family background, differential treatment due to gender and race, and different access to public services due to place of birth.

To complete the list of determinants, let $z$ denote the additional set of socially-determined but unobserved circumstances and $a$, the set of non-socially determined circumstances. Finally, let $w$ denote measures of efforts. Hence, $y = \psi(x, z, a, w)$. Additionally, assume that either the factors $(z, a, w)$ are not observed or we have decided not to control for them. Hence, from a practical point of view they can be treated as unobservable. To simplify, assume the relation determining the outcome is log-additive. Hence,

$$\ln(y) = g(x) + h_1(z) + h_2(a) + h_3(w)$$

Given that $y$ and $x$ are observed we could estimate, for instance, the log-mean regression of $y$ on $x$, i.e., we could estimate $\mu(x) \equiv E(\ln(y)|x)$. Unless the excluded variables $(z, a, w)$ are mean-independent of the observed socially-determined circumstances, $x$, it will not be, in general, true that $g(x) = \alpha + \mu(x)$. If they were mean-independent, then, of course, this relation would be valid with

$$\alpha = -(E(h_1(z)) + E(h_2(a)) + E(h_3(w)))$$

In general, however,

$$\mu(x) = g(x) + E(h_1(z)|x) + E(h_2(a)|x) + E(h_3(w)|x) \quad (1)$$

Even though it would be nice to treat all choices as effort and all exogenous variables as circumstances, this separation is not always possible. Consider a worker who chooses to live in a nearby favela and to be self-employed, as opposed to live in a distant community and work as an employee in a factory. Let us assume this choice is rational and improve his welfare. In addition, suppose that teachers discriminate children from favela and employer discriminate workers from favela. As a result of the discrimination this worker may prefer to be an informal self-employed and his children may evade school earlier. In this case, the advantages of proximity with a better labor market dominate the discrimination effects. For this family it is better to be discriminated but to live nearby the labor market, than to avoid being discriminated by working in a distant factory as a low-paying employee. Nevertheless, it remains true that this family standard of living would be much better if there were no discrimination against favela. As it is always the case, discrimination is a factor leading to equally hard-working people to achieve different outcomes. It should, therefore, be considered a source of inequality of opportunity, even when families choose the situation where they are discriminated.
In our practical applications we are going to use the inequality in $\mu(x)$ to measure the dependence of the outcome on socially-determined circumstances and to construct an index of inequality in socially-determined opportunities. As expression (1) reveals, this index will be also capturing a host of other effects in addition to the impact of observed socially-determined circumstances on outcome. In the next section we aim to organize and illustrate these potential departures. The reader only interested in the computation of the inequality of opportunity index could skip the next section, since its contents matter only for the interpretation of the measure we propose.

1.4. WHY WOULD OUTCOMES AND OBSERVED SOCIALLY DETERMINED CIRCUMSTANCES BE CORRELATED?

As expression (1) aims to clarify, in addition to their direct causal link, there are at least three reasons for an outcome to be correlated with observed socially-determined circumstances. Their formal origins are well described in expression (1). In this section we aim to illustrate their substantive content.

The association between observed and unobserved socially-determined circumstances $E(h_i(z)|x) \neq E(h_i(z))$

To the extent that observed and unobserved socially-determined circumstances are correlated, $\mu(x)$ will also capture at least part of the impact of unobserved socially-determined circumstances. For instance, to the extent that living in a favela is correlated with race, the fraction of the inequality of opportunity we are going to assign to race will bias its own real effect. In fact, even when teachers do not discriminate by race, but do discriminate residents from favela, we are going to capture some inequality in educational opportunity by race. Hence, our indicator has difficulties in isolating the contribution of each observed circumstance.

The association between observed socially-determined circumstances and intrinsic circumstances $E(h_2(a)|x) \neq E(h_2(a))$

To the extent that intrinsic circumstances and observed socially-determined circumstances are correlated, $\mu(x)$ will also capture at least part of the impact of intrinsic circumstances. For instance, suppose that parents’ talents are correlated with their children talents and own talent is correlated with own education. In this case, some of the inequality of opportunity we assign to differences in parents’ education may overestimate their contribution. In fact, even when parents’ education has no effect on their children education, we are still going to capture some inequality of opportunity associated to parents’ education due to the inheritance of talents. Again, our indicator has difficulties in isolating the real contribution of a specific circumstance.
The association between effort and observed socially-determined circumstances
\[ E(h_1(w)|x) \neq E(h_3(w)) \]

Suppose that differences by race in educational effort are a reaction to discrimination in school. In this case, the inequality of opportunity we assign to race includes not only the direct effect of discrimination in school, but also its indirect effect on effort. Our indicator will not be comparing outcomes between students who are putting the same amount of effort. Our indicator will consider all differences by race including those related to effort.

Even more worrisome is the possibility that differences in effort – which we are including in our measure of inequality of opportunity – may be due to the anticipation of future discrimination in the labor market. Hence, we could estimate a sizeable degree in educational opportunity related to race, even when schools are a perfect non-discriminatory institution. This example illustrates that our measure may have severe difficulties in identifying where the inequality of opportunity really occurs.

Furthermore, since differences in public policy (for instance, any form of positive discrimination) may influence effort, gender differences in work effort could be, for instance, related to the generosity of maternity benefits. In this case, we are going to consider as inequality of opportunity differences in effort generated by discriminatory social policy.

The association between observed socially-determined circumstances, risks and other exogenous random events

Some groups may choose greater risks or they may be forced to bear greater risks. In both cases, outcomes may be related to group membership, even though there are no other forms of discrimination. For instance, black children may have worse performance at school because they have poorer health due to uneven access to basic preventive health care. Moreover, due to their greater poverty, blacks may choose more hazardous working activities and as a consequence spent a larger fraction of their life without working, due to poor health conditions. In the first case, we are going to estimate a degree of inequality in educational opportunity that is unrelated to schools. In the second case, we are going to estimate a degree of inequality in working opportunity that is totally related to differences in endogenous choice.

Summary

As all these situations clearly emphasize, we have to be very careful in interpreting the measure of inequality of opportunity we are going to estimate. It may be capturing differences in effort and luck stochastically associated to circumstances. It can also be ascribing differential treatment in the labor market or in the access to health to inequality in educational opportunity. Finally, it may be attributing differences to socially-determined inequality of opportunity due to intrinsic circumstances like talents.
1.5. The direct and the residual approaches

If the inequality of opportunity is defined as the inequality in average outcome between groups defined in terms of their common circumstances, we could estimate it based on two types of counterfactuals. On one hand, we could define it as the outcome inequality that would remain if all the inequality within groups with the same circumstances were eliminated. On the other hand, we could define it as the difference between the overall outcome inequality and the level that would remain if all the inequality in means between groups with the same circumstances were eliminated. In the sequence, we define precisely each of these two approaches and identify when they are equal.

*Inequality of opportunities as the between component of the inequality in outcomes*

To isolate the outcome inequality between groups it is enough to eliminate the inequality within groups and re-compute the overall inequality. More precisely, we change the outcome for every person to equal the average of her circumstance group. More precisely, we make

\[ y_b(x) = E(y|x = x(\omega)) = \lambda(x(\omega)) \]

where \( x = x(\omega) \) represents the set of circumstances of a specific group. In this case, if \( I(y) \) is a proper measure of inequality in the distribution of \( y \), then the inequality of opportunity \( IO_b(y|x) \) will be given by \( IO_b(y|x) = I(y_b) \). Notice that this measure of inequality of opportunity does not require \( y \) to be observed or the overall inequality \( I(y) \) to be measured, as long as the function \( \lambda \) is known and the socially-constructed circumstances observed.

Let’s work the precise expression of this indicator for the Theil-L inequality measure. To begin with, notice that by its definition

\[ L(y) = \ln(E(y)) - E(\ln(y)) \]

Hence,

\[ IO_b(y|x) = L(y_b) = \ln(\lambda(x)) - E(\ln(\lambda(x))) \]

The relation between the overall inequality and the inequality in opportunities can be perceived by the expressions

\[ L(y) = E(\ln(E(y|x))) - E(\ln(y|x)) + \ln(E(E(y|x))) - E(\ln(E(y|x))) \]

or
\( L(y) = E(\ln(\lambda(x)) - E(\ln(y|x))) + \ln(E(\lambda(x))) - E(\ln(\lambda(x))) \)

Since,
\( L(y_b) = \ln(\lambda(x)) - E(\ln(\lambda(x))) \)

Hence,
\( L(y) = E(\ln(\lambda(x)) - E(\ln(y|x))) + L(y_b) \)

**Inequality of opportunities as the difference between the overall outcome inequality and the within component**

In this approach we estimate the inequality in opportunity by how much the outcome inequality would decline if the inequality between circumstance groups were eliminated. To estimate the inequality of opportunity in this case, it is therefore necessary to isolate the within component of the outcome inequality, i.e., the inequality that would remain if all inequality between circumstance groups were eliminated. To isolate the inequality within groups, we should change the outcome for every person proportionally to the gap between the average for the group and the overall average. More precisely, we make

\[ y_d(\omega) = \frac{\mu_y(\omega)}{\lambda(x(\omega))} \]

In this case if \( I(y) \) is a proper measure of inequality in the distribution of \( y \), then the inequality in opportunity \( IO_d(y|x) \) will be given by \( IO_d(y|x) = I(y) - I(y_a) \).

Next, we obtain the precise expression of this indicator for the Theil-L. To begin with, notice that by its definition
\( L(y) = \ln(E(y)) - E(\ln(y)) \)

Hence,
\( L(y_d) = \ln \left( E \left( \frac{\mu_y}{\lambda(x)} \right) \right) - E \left( \ln \left( \frac{\mu_y}{\lambda(x)} \right) \right) = E(\ln(\lambda(x))) - E(\ln(y)) \)

The relation between the overall inequality and the within inequality can be expressed as
\( L(y) = (E(\ln(\lambda(x))) - E(\ln(y))) + L(y_b) \)

Hence,
\[ L(y) = L(y_d) + L(y_b) \]

Therefore

\[ IO_d(y|x) = L(y) - L(y_d) = L(y_b) = IO_b(y|x) \]

2. **MEASURING INEQUALITY OF OPPORTUNITY FOR DISCRETE OUTCOMES**

If the inequality of opportunity is defined as the inequality in average outcome between groups defined in terms of their common circumstances, we could estimate it based on two types of counterfactuals. As explained in the previous section, we could define it as the outcome inequality that would remain if all the inequality within groups with the same circumstances were eliminated: the direct approach. On the other hand, we could define it as the difference between the overall outcome inequality and the level that would remain if all the inequality in means between groups with the same circumstances were eliminated: the indirect approach. In the case of discrete outcomes the direct approach has important advantages. For this reason we based our development of a measure of inequality of opportunity for discrete outcomes on this approach.

2.1. **THE ADVANTAGES OF THE DIRECT APPROACH FOR MEASURING INEQUALITY OF OPPORTUNITIES FOR DISCRETE OUTCOMES**

As already mentioned, the direct approach requires only the knowledge of the regression function, whereas the residual approach would need information on outcomes and regression residuals. When outcomes are completely observed, the two approaches convey equivalent information. However, when outcomes are only partially observed the direct approach may have great practical advantages. In this section we clarify these advantages.

As before, assume that, \( y = \lambda(x) + \varepsilon \). But now assume that we only observe if the outcome is smaller or greater than \( \eta \), let \( d \) be an indicator of this event. Hence, \( d = 1 \iff y > \eta \) and \( d = 0 \iff y \leq \eta \). The only observed variables are the indicator \( d \) and the socially-determined circumstances \( x \).

Since, \( \mu(x) \equiv E(y|x) = 0 \). Let strengthen this condition, assuming that \( \varepsilon \) is stochastically independent of \( x \) and have a known distribution, \( F_\varepsilon \). Hence,

\[
E(d|x) = P(d = 1|x) = P(y > \eta|x) = P(\varepsilon > \eta - \mu(x)) = 1 - F_\varepsilon(\eta - \mu(x))
\]
Since the pair \((d, x)\) is observed we can estimate the regression function \(E(d|x)\). To the extent that \(F_{\varepsilon}\) is known, we could obtain \(\mu(x)\), up to a constant, without fully observing \(y\), via

\[
\mu(x) = \eta - F_{\varepsilon}^{-1}(1 - E(d|x))
\]

and proceed to construct a measure of inequality of opportunity as described in the previous section.

It must, however, be clarified that given all these hypotheses the residual approach could also be pursued. In fact, the distribution of the residuals we know by hypothesis. Therefore, once we have estimated the regression function, we could obtain the outcome distribution, since by hypothesis \(\varepsilon\) is stochastically independent of \(x\). It is still true that we can not recover the outcome values for each person, but its distribution is completely determined combining the observed data with the hypotheses we have made.

To assume the distribution of \(\varepsilon\) is known may be a very high price to be paid to construct an indicator of equality of opportunities. Mainly, because we can avoid this hypothesis if we construct an inequality of opportunity measure entirely based on \(E(d|x)\). Let us denote \(E(d|x)\) by \(p(x)\), i.e., \(p(x) = E(d|x)\). If, instead of insisting in measuring the inequality of opportunity from the inequality in \(\mu(x)\) we accept to measure it from the inequality in \(p(x)\), we could obtain a measure of inequality in opportunity without having to assume a distribution for effort.

We have, nevertheless to keep the hypothesis that effort is stochastic independent of circumstances, otherwise \(E(d|x)\) would not be a function of \(x\) only through \(\mu(x)\), but also due to the stochastic relation between effort and circumstances. In other words, for \(E(d|x) = 1 - F_{\varepsilon}(\eta - \mu(x))\) we require \(\varepsilon\) and \(x\) to be stochastically independent, otherwise we would have \(E(d|x) = 1 - F_{\varepsilon}(\eta - \mu(x)|x)\), and there would be two sources of dependence on \(x\) built in this regression: one related to the direct impact of circumstance on outcome, the source of inequality of opportunity we want to measure, and another due to the nuisance stochastic dependence of effort on circumstance.

**Measuring differences between distributions**

Given estimates for the regression function \(p(x)\) and for the distribution of \(x\), many measures of inequality could be easily computed. All of them would be measures of the inequality of opportunity. However, the binary nature of \(d\) and the fact that \(p(x)\) is a conditional probability and so bounded between 0 and 1 permit and may also require specialized treatment. Let \(\overline{p} = E(d) = P(d = 1)\).
We begin noticing that, as a consequence of the binary nature of \( d \), mean-independence is equivalent to full stochastic independence. Hence,

\[
p(x) = \overline{p} \iff F_{Y=1} = F_{Y=0}
\]

In other words, the probability of a low outcome is independent on the circumstances, if and only if, the distribution of circumstances among those with low outcomes is identical to the distribution of circumstances among those with high outcome. Therefore, we could measure the degree of inequality of opportunity by computing the distance from the distribution of circumstances for those with low outcomes to the distribution of circumstances for those with high outcomes. One advantage of this procedure is that we could select any measure from the large literature on measuring distances between two distributions.

Two measures are widely used. In sociology and demography, the Dissimilarity Index is the most commonly used. In statistics, the Kullback-Leibler measure is probably preferred. We are going to concentrate our attention on the dissimilarity index due to its immediate, intuitive and simple interpretation.

Let us consider a situation where circumstances have a discrete distribution and \( \{x_1, \ldots, x_m\} \) denotes the set of all possible values. Then, the dissimilarity between the distribution of circumstances among those with high outcomes, \( f_i(x_k) = P(x = x_k | y > \eta) \) and the distribution of circumstances among those with low outcomes, \( f_0(x_k) = P(x = x_k | y \leq \eta) \) is defined as

\[
D = \Lambda \sum_{k=1}^{m} |f_i(x_k) - f_0(x_k)|
\]

where, \( \Lambda = \frac{1 - \overline{p}}{2} \) and as defined before \( \overline{p} = P(y > \eta) = P(d = 1) \). The Dissimilarity Index is, therefore, proportional to the absolute distance between the two distributions. The proportional factor \( \Lambda \) plays an important role in providing a simple interpretation for this index. Before considering its interpretation, notice that it could be alternatively expressed as the absolute distance between the distribution of circumstances among those with high outcomes, \( f_i(x_k) = P(x = x_k | y > \eta) \) and the overall distribution of circumstances, \( f(x_k) = P(x = x_k) \). In fact,

\[\text{KL} = \sum_{k=1}^{m} f_i(x_k) \cdot \ln \left( \frac{f_i(x_k)}{f(x_k)} \right)\]

An alternative to the dissimilarity index is the Kullback-Leibler measure. As a measure of the distance between the distribution of circumstance among the high outcome persons and the overall distribution of circumstances this measure would be given by
Moreover, since \( f(x_k).p(x_k) = f_1(x_k).\bar{p} \), it follows that the Dissimilarity Index could also be expressed as

\[
D = \Gamma \sum_{k=1}^{m} |p(x_k) - \bar{p}|f(x_k)
\]

where, \( \Gamma = \frac{1}{2\bar{p}} \). Hence, the Dissimilarity Index is proportional to the area between the mean-regression function, \( p(x) \), and the overall average \( \bar{p} \) (see Figure 2).

Figure 2: Graphical representation for Dissimilarity Index

To see the usefulness of the proportional factors, \( \Lambda \) and \( \Gamma \), consider a finite population with \( N \) members. Hence, \( L = N.\bar{p} \) would be the total number of persons with high outcome, and \( L(x_k) = N.f(x_k).p(x_k) \) the corresponding number among those with circumstances equal to \( x_k \). Moreover, notice that \( \bar{L}(x_k) = N.f(x_k).\bar{p} \) would be the expected number of persons with high outcome if equal opportunity prevails. Hence, a natural measure of inequality in opportunity would be
\[ IO = \frac{1}{2L} \sum_{k=1}^{m} |L(x_k) - \bar{L}(x_k)| \]

Notice that \( IO = D \). Moreover, \( IO \) represents the minimum fraction of the total number of persons with high outcome, \( L \), that needs to be redistributed across circumstance groups in order to achieve equal opportunity, i.e., to ensure equal proportion to low outcome persons in all circumstance groups, \( p(x_i) = \bar{p} \).

### 2.2. Computing the Dissimilarity Index

Assume that one has access to a random sample of the population with information on whether person \( i \) had or not access to a given opportunity \( (I_i = 1 \) if that person had access and \( I_i = 0 \) otherwise) and a vector of variables indicating his/her circumstances, \( x_i = (x_{i1}, \ldots, x_{im}) \).

Given this information, one need to follow three steps to estimate the inequality of opportunity index

\[ D = \frac{E|P(I = 1|x) - P(I = 1)|}{2P(I = 1)} \]

Before we proceed, however, it is worth noticing that since,

\[ P(I = 1) = E(P(I = 1|x)) \]

we can rewrite \( D \) as

\[ D = \frac{E|P(I = 1|x) - E(P(I = 1|x))|}{2E(P(I = 1|x))} \]

which is actually the expression we mimic in order to estimate \( D \). This expression also indicates the central role of group specific coverage rates, \( P(I = 1|x) \), in estimating \( D \).

These conditional probabilities could be estimated through a variety of parametric, nonparametric or semi-parametric procedures. One could impose separability restrictions or consider interactions. In all case the three step procedure we describe in this note would apply.
Since we were planning to apply this procedure to all Latin America and Caribbean countries with available data and for several points in time we privileged a standard specification that could be feasibly applied to all countries at all times. Our choice was a separable logistic model. Hence, our first step in estimating $D$ was to fit the following separable logistic regression

$$
\ln \left( \frac{P(I = 1|x_1,\ldots,x_m)}{1 - P(I = 1|x_1,\ldots,x_m)} \right) = \sum_{k=1}^{m} h_k(x_k)
$$

where $x_k$ denotes a vector of variables representing the $k$-dimension of circumstances, hence, $x = (x_1,\ldots,x_m)$. The dimensions consider include parents education, family per capita income, gender, number of siblings, family structure (number of siblings, single-parent household) and area of residence (urban versus rural). The functions $\{h_k\}$ were chosen accordingly to the needs of each dimension: quadratic on education, logarithmic on income, nonparametric (dummies) on age and the other dimensions. In all cases all functions end up being linear in the parameters, so that, $h_k(x_k) = x_k \beta_k$. A complete specification of this logistic regression is presented in Table 1. From the estimation of this logistic regression one obtains estimates of the parameters $\{\beta_k\}$ that will be denoted by $\{\hat{\beta}_k\}$.

### Table 1: Specification of separable logistic regression function

<table>
<thead>
<tr>
<th>Dimensions of circumstances</th>
<th>Specification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>free (dummy)</td>
</tr>
<tr>
<td>Parents education</td>
<td>quadratic</td>
</tr>
<tr>
<td>Per capita income</td>
<td>logarithmic</td>
</tr>
<tr>
<td>Number of siblings</td>
<td>linear</td>
</tr>
<tr>
<td>Presence of parents</td>
<td>free (dummy)</td>
</tr>
<tr>
<td>Area of residence (urban versus rural)</td>
<td>free (dummy)</td>
</tr>
</tbody>
</table>

---

7 In the case of education age was also a variable used to predict the probability of completing each grade.

8 The results for these logistic regressions are shown in Tables A.3.1 to A.3.6 at the end of this paper.
As our second step, given this coefficient estimates we obtain for each individual in our universe his/her predicted probability of access to the opportunity in consideration, i.e., for each individual $i$ we compute

$$\hat{p}_i = \frac{\exp\left(\hat{\beta}_0 + \sum_{k=1}^{m} x_{ki} \hat{\beta}_k\right)}{1 + \exp\left(\hat{\beta}_0 + \sum_{k=1}^{m} x_{ki} \hat{\beta}_k\right)}$$

Next, as a third and final step we compute

$$\bar{p} = \sum_{i=1}^{n} w_i \hat{p}_i$$

and

$$\hat{D} = \frac{1}{2\bar{p}} \sum_{i=1}^{n} w_i |\hat{p}_i - \bar{p}|$$

where $w_i = \frac{1}{n}$ or some sampling weights.  

Since, almost surely, $\lim_{n \to \infty} (\bar{p}) = P(I = 1)$ and, under the assumptions that (i) the regression has been correctly specified and (ii) its coefficients consistently estimated, also

$$\lim_{n \to \infty} \left(\sum_{i=1}^{n} w_i |\hat{p}_i - \bar{p}|\right) = E[P(I = 1|x) - P(I = 1)]$$

almost surely. It follows then that $\lim_{n \to \infty} (\hat{D}) = D$ almost surely. In other words $\hat{D}$ is a consistent estimator of $D$.

Since $\hat{D}$ is a ratio of two linear functions of $\hat{p}_i$, it can be shown that the asymptotic variance of $\hat{D}$, $\sigma_D^2$, is given by $\sigma_D^2 = \Gamma \Omega \Gamma'$, where

$$\Gamma = \frac{1}{\bar{p}^2} \left[\left(\sum_{i \in L} w_i \hat{p}_i\right)\left(\sum_{i \in L} w_i \hat{p}_i(1 - \hat{p}_i)x_i\right) - \left(\sum_{i \in U} w_i \hat{p}_i\right)\left(\sum_{i \in L} w_i \hat{p}_i(1 - \hat{p}_i)x_i\right)\right]$$

---

9 The Dissimilarity index for the 19 selected Latin American countries are shown in Table A.3.7 at the end of this paper.
\( \Omega_{\beta} \) is the asymptotic variance matrix for \( \hat{\beta} \) and \( L \) denotes the set of all individuals with predicted probability of access below average and \( U \) the complement. So, \( L = \{ i : \hat{p}_i \leq \bar{p} \} \) and \( U = \{ i : \hat{p}_i > \bar{p} \} \)

We conclude considering two practical problems. First, it must be emphasized that the entire procedure is consistent if we keep or if we drop the non-significant parameters. Asymptotically the choice is irrelevant. We, however, have decided for keeping the non-significant coefficients. Our choice was based on a couple of considerations. Firstly, we have to recognize that dropping a coefficient will always reduce the measured level of inequality. Since the presence of non-significant coefficients depends as much on their magnitude as on sample size, dropping them would favor smaller sample size countries. Secondly, by the same reason that we do not revert to 2 all coefficients that are not significantly different from 2, we have no reason to give zero a special treatment.

A second practical question is the treatment to be given to coefficients with sign contrary our expectations. For instance, what should we do when rural areas are estimated to have greater access than the corresponding urban areas? We opted for keeping all the coefficients at their estimated values. Again our decision was based on a couple of reasons. First, because respecting the data could prove informative. Today, throughout Latin America girls are performing better than boys in school, so the male dummy receives a consistent negative sign. We are all now quite familiar with these findings, so we would not consider their sign incorrect. But the first econometricians encountering this finding a decade ago could consider it unlikely and discard the information. Secondly, most of our intuition and reasons for discarding unexpected signs may be based on univariate comparisons. We all expect the access to opportunities to be worse in rural areas, but are we absolute sure about this contrast after we have extensively control for parents’ education, family income and composition?

2.3. Properties of the Dissimilarity Index

Range of the Inequality of Opportunity Index

Consider a situation in which we have \( m \) groups. Let \( p_i \) be the proportion of children in group \( i \) with access to a basic social service and let \( \beta_i \) be the proportion of children in this group. In this case our new dissimilarity index will be given by

\[
D = \frac{1}{2\bar{p}} \sum_{i=1}^{m} \beta_i |p_i - \bar{p}|
\]

where
\( \bar{p} = \sum_{i=1}^{m} \beta_i p_i \)

**Proposition 1:** \( D \geq 0 \)

In fact, since \( |p_i - \bar{p}| \geq 0 \), it follows that \( \sum_{i=1}^{m} \beta_i |p_i - \bar{p}| \geq 0 \). Hence, \( D \geq 0 \).

**Proposition 2:** \( D \leq (1 - \bar{p}) \)

Let us, without any loss of generality, order the groups increasingly according to their values of \( p_i \). Let \( k \) be such that \( p_k < \bar{p} \) and \( p_{k+1} \geq \bar{p} \). Hence, \( k \) is the number of groups with probability of access below average. In this case, the numerator of the dissimilarity index can be rewritten as:

\[
\sum_{i=1}^{m} \beta_i |p_i - \bar{p}| = \sum_{i=1}^{k} \beta_i (\bar{p} - p_i) + \sum_{i=k+1}^{m} \beta_i (p_i - \bar{p}) = \bar{p} \left( \sum_{i=1}^{k} \beta_i - \sum_{i=k+1}^{m} \beta_i \right) - \left( \sum_{i=1}^{k} \beta_i p_i - \sum_{i=k+1}^{m} \beta_i p_i \right)
\]

hence,

\[
\sum_{i=1}^{m} \beta_i |p_i - \bar{p}| = \bar{p} \left( \sum_{i=1}^{m} \beta_i - 2 \sum_{i=k+1}^{m} \beta_i \right) - \left( \sum_{i=1}^{m} \beta_i p_i - 2 \sum_{i=k+1}^{m} \beta_i p_i \right) = 2 \left( \sum_{i=k+1}^{m} \beta_i p_i - \bar{p} \sum_{i=k+1}^{m} \beta_i \right)
\]

therefore,

\[
\sum_{i=1}^{m} \beta_i |p_i - \bar{p}| = 2 \left( \sum_{i=k+1}^{m} \beta_i p_i - \bar{p} \sum_{i=k+1}^{m} \beta_i \right) = 2 \left( \bar{p} - \left( \sum_{i=1}^{k} \beta_i p_i + \bar{p} \sum_{i=k+1}^{m} \beta_i \right) \right)
\]

and

\[
\sum_{i=1}^{k} \beta_i p_i + \bar{p} \sum_{i=k+1}^{m} \beta_i > \bar{p} \sum_{i=1}^{k} \beta_i p_i + \bar{p} \sum_{i=k+1}^{m} \beta_i p_i = \bar{p}^2
\]

Therefore,

\[
\sum_{i=1}^{m} \beta_i |p_i - \bar{p}| \leq 2\bar{p}(1 - \bar{p})
\]

implying that

\[
D = \frac{1}{2\bar{p}} \sum_{i=1}^{m} \beta_i |p_i - \bar{p}| \leq 1 - \bar{p}
\]
Next let us now consider two extreme situations:

**Case A:** \( p_i = \overline{p} \) for all \( i = 1, \ldots, m \).

**Case B:** \( p_i = 0 \) for all \( i = 1, \ldots, m-1 \) and \( p_m = 1 \).

**Proposition 3:** In Case A, \( D = 0 \)

In fact, in this case, \( |p_i - \overline{p}| = 0 \) for all \( i \) and consequently, \( D = 0 \).

**Proposition 4:** In Case B, \( D = 1 - \beta_m = 1 - \overline{p} \). Therefore, \( D \uparrow 1 \) as \( \beta_m \downarrow 0 \).

Since in thus case \( p_i = 0 \) for all \( i = 1, \ldots, m-1 \), \( \overline{p} = \beta_m p_m \). Consequently, for all \( i = 1, \ldots, m-1 \), \( |p_i - \overline{p}| = \overline{p} = \beta_m p_m \). Hence,

\[
\sum_{i=1}^{m-1} \beta_i |p_i - \overline{p}| = \sum_{i=1}^{m-1} \beta_i \beta_m p_m = (1 - \beta_m) \beta_m p_m
\]

In addition, since

\[
|p_m - \overline{p}| = (1 - \beta_m) p_m
\]

\[
\sum_{i=1}^{m} \beta_i |p_i - \overline{p}| = \sum_{i=1}^{m-1} \beta_i |p_i - \overline{p}| + \beta_m |p_m - \overline{p}| = (1 - \beta_m) \beta_m p_m + (1 - \beta_m) \beta_m p_m = 2(1 - \beta_m) \overline{p}
\]

Hence,

\[
D = \frac{1}{2 \overline{p}} \sum_{i=1}^{m} \beta_i |p_i - \overline{p}| = \frac{1}{2 \overline{p}} 2(1 - \beta_m) \overline{p} = (1 - \beta_m)
\]

In sum, \( D = 1 - \beta_m \). As a consequence, \( D \uparrow 1 \) as \( \beta_m \downarrow 0 \).

**Sensitivity of Dissimilarity Index to variations in the coverage rate**

In principle, any *inequality* of opportunity index to deserve this denomination must be insensitive to a *balanced* increase in the availability of opportunities. In fact, new opportunities that are allocated using the same criteria as the pre-existing ones should not have an effect on the level of inequality. That a proper measure of inequality of opportunity must have this property is clear, the difficulty is how to define the concept of *balanced increase* in availability of opportunities.

In measuring income inequality, balanced growth is easy to define. To obtain balanced growth one has just to distribute the new income in the same way previous income was distributed. If a person holds \( \delta \% \) of the original total income, he/she must receive \( \delta \% \) of
the additional income. As a consequence, he/she will remain with $\delta\%$ of the new total income. In this case, every and each person will receive some extra income\(^{10}\).

Some specific opportunities, either one has or does not have, as opposed to income where there may be large variations among those who have. So, an increase in the availability of opportunities necessarily means to provide access for those who did not have access previously, leaving those who already had access unchanged. In fact, when we are dealing with a specific opportunity, we can not increase the opportunity of those who already had access to it. In this case the notion of balanced increase must naturally be more complex.

To construct a balanced increase in this situation, we have to change the unity of analysis from individuals to circumstance groups. This is not a limitation, since measures of the inequality of opportunity should be insensitive to the distribution of opportunities within circumstance groups. So, we restrict our attention to the average access of each group. Accordingly, a balanced increase would be a situation in which new opportunities are assigned to circumstances groups in the same way as the pre-existing ones were in the past.

It seems, however, that we have just changed labels. We have defined balanced increase based on the assignment of opportunities to circumstance groups in the same way as the pre-existing opportunities were originally assigned. But what we mean by in the same way? A natural interpretation is to assume that for a balanced increase to occur one need the distribution of the new opportunities to be distributed among circumstance groups in the same way as the pre-existing distributions are. In this case, the distribution of opportunities across circumstance groups would remain unchanged.

In this case, due to the increase in the availability of opportunities both the overall coverage, $\bar{p}$, and the group specific access probabilities, $\{p_j\}$, would increase. Given the way the new opportunities are distributed, all specific probabilities increase proportionally to the overall coverage rate. Hence, if $p_j^*$ denotes the new group $i$ specific access probability, then $p_j^* = (1 + \lambda)p_j$ for all $j$ and also $\bar{p} = (1 + \lambda)\bar{p}$. So, $\frac{p_j}{\bar{p}} = \frac{p_j^*}{\bar{p}^*}$ since neither the proportion of the population in each group, $\beta_j$, neither the proportion of opportunities allocated to each group, $\gamma_j$, change as a result of this balanced increase in opportunities (remember that $\frac{p_j}{\bar{p}} = \frac{\gamma_j}{\beta_j}$).

As long as we measure the inequality of opportunity using $D$, defined as before

\(^{10}\) An exception will be those that originally had no income. To keep the balance, they will receive nothing.
\[ D = \frac{1}{2} \sum_{j=1}^{m} \beta_j \left| p_j - \bar{p} \right| = \frac{1}{2} \sum_{j=1}^{m} \left| \gamma_j - \beta_j \right| \]

and the inequality of opportunity will be insensitive to this type of balanced increase in opportunities.

Traditionally, however, the dissimilarity index has been defined slightly differently. In most cases is defined in terms of the lack of access, i.e.,

\[ D_t = \frac{1}{2(1-\bar{p})} \sum_{j=1}^{m} \beta_j \left| (1-p_j) - (1-\bar{p}) \right| = \frac{1}{2(1-\bar{p})} \sum_{j=1}^{m} \beta_j \left| p_j - \bar{p} \right| \]

In this case, the interpretation changes slightly. As mentioned before, \( D \) is the proportion of all opportunities that need to be rearranged to ensure equal access for all groups. In other words, is the amount of opportunities that need to be rearranged as a proportion of the number of individuals who already have access. The more traditional measure, \( D_t \), keeps the same numerator, the amount of opportunities that need to be rearranged for equal opportunity to prevail, but changes the denominator to become the number of individuals without access. So, it becomes the amount of opportunities that need to be rearranged as a proportion of individuals without access to this opportunity.

In short, \( D \) centers the attention on access and \( D_t \) on the lack of access. In fact, \( D = \frac{1}{2} \sum_{j=1}^{m} \left| \gamma_j - \beta_j \right| \). Hence, it measures the distance between the distribution of the overall population, \( \{\beta_j\} \), and the distribution of individuals with access to the opportunity of interest, \( \{\gamma_j\} \), across groups. On the other hand, \( D_t = \frac{1}{2} \sum_{j=1}^{m} \left| \eta_j - \beta_j \right| \), where \( \eta_j \) denotes the proportion of individuals without access to the opportunity in consideration who belong to circumstance group \( j \). So it measures the distance between the distribution of the overall population, \( \{\beta_j\} \), and the distribution of individuals without access across opportunity groups, \( \{\eta_j\} \).

In fact, a more symmetric measure, \( D_2 \), would be the distance between the distribution of the individuals with access, \( \{\gamma_j\} \), and the distribution of individuals without access across opportunity groups, \( \{\eta_j\} \). In this case,

\[ D_2 = \frac{1}{2} \sum_{j=1}^{m} \left| \gamma_j - \eta_j \right| \]

It can be shown that in this case \( D_2 \) can also be written as
Above we demonstrate that $D$ is insensitive to a *balanced* increase in the availability of opportunities, at least as we have defined it above. $D_1$ and $D_2$, however, do not share this property\(^{11}\). Both measures would increase, indicating that a balanced increase in the availability of opportunities (one that would keep the distribution of opportunities across groups unchanged) would *increase* the inequality of opportunity as measured by these two indices. This undesirable property is one of the main reasons why we opted to measure the inequality of opportunity using $D$.

Although our concept of balanced increase in opportunity is quite natural, it has at least one important practical limitation worth mentioning. In fact, it does not respect the restriction that each individual can have at most one of the specific opportunity we are considering (in fact, either one has or does not have access to school), so that we always must have $p_j \leq 1$. If the coverage is already very high in certain groups, it might not exist a way of allocating the new opportunities as the old ones were allocated. Suppose we have two groups and all the original opportunities were allocated just to one of the groups ensuring full coverage for this group. In this case, new opportunities cannot be allocated to the advantage group because their coverage rate is already 100%. Moreover, the new opportunities cannot be allocated either to the disadvantage group without disturbing the distribution of opportunities across groups. So, using this concept, no balanced increase in opportunity is possible in this situation. Given this drawback, one might be compel to review the concept. However, it does not seem to exist a simple alternative without its own limitations. Let us consider a couple of them for further analysis.

One possibility would be to consider as *balanced* an increase in opportunities in which the new opportunities are distributed at random. This would be equivalent to distribute opportunities across groups according to their population size. In this case the access probability of all circumstance groups will increase by the same amount. So, while according to our previous concept of *balanced* increase in opportunities $\Delta p_i = \Delta p + \lambda$, under this alternative concept $p_i^* = p_i + \lambda$.

In this case, the absolute number of opportunities that need to be reallocated, $\frac{N}{2} \sum_{j=1}^{m} \beta_j |p_j - \bar{p}|$, would not be influenced by the any balanced increase. As a consequence, following a balanced increase in opportunity of this type $D$ would decline, $D_1$ will increase and $D_2$ will decrease if $\bar{p} < 1/2$ and increase if $\bar{p} > 1/2$. So, when the

\[^{11}\text{The proof is simple. First, notice that } D_2 = D + D_1. \text{ Since } D \text{ is insensitive to a balanced increase in opportunities, the sensitivity of } D_1 \text{ and } D_2 \text{ must be the same. Next notice that } D_2 = \frac{1}{1-\bar{p}} D. \text{ Hence, any increase in } \bar{p} \text{ not affecting } D \text{ would increase } D_2 \text{ and so } D_1.\]
majority of the population already has access to the opportunity in consideration, out of the three measures only $D$ will decline as a consequence of such balance increase in opportunity.

It should be noticed that this concept share the same limitation as the previous one. The restriction that no individual can have more than one opportunity implies that no balanced increase may be possible in some situations.

A second and more elaborate alternative would be to consider balanced proportional increase in the odds ratio of all circumstance groups. In this case, a balanced increase in opportunities occur when

$$\frac{p_j^*}{1-p_j^*} = \left(1 + \lambda \right) \frac{p_j}{1-p_j}$$

which is equivalent to

$$p_j^* = \frac{1 + \lambda}{1 + \lambda p_j} \cdot p_j$$

So, in this case the proportional increase in specific coverage will vary accordingly the baseline situation, with the proportional increase being greater for the more disadvantage groups. For groups with original coverage closer to zero the coverage will increase by $\lambda \%$, whereas for groups with original coverage closer to universal coverage will not be affected.

It can be shown that also in this case the inequality as measure by $D$ would always decline as a consequence of a balanced increase in opportunity of this type. The other alternative measures, $D_1$ and $D_2$, may decline or even increase. In any case, out of the three measures $D$ would be the one experiencing the greatest decline.

In sum, we consider three alternative measures for the inequality of opportunity, $D$, $D_1$ and $D_2$, and their response to three notions of a balanced increase in the availability of opportunities. Table 1 summarizes the main results. Ideally, we would like these measures to be insensitive to a balanced increase in opportunities. The fact that $D_1$ would always increase and $D_2$ would increase if $\bar{p} > 1/2$ (the most common case) if one distributes the new opportunities at random, lead us to strongly prefer the measure $D$. $D$ would always decline if the newly available opportunities were distributed at random, The preference for the measure $D$ also comes from the fact that it is insensitive to an increase in opportunities whenever the new opportunities are distributed across groups accordingly the distribution of the pre-existing opportunities. In this case, both $D_1$ and $D_2$ would increase (see Table 2). If the concept of a balanced increase in opportunities is taken to represent a proportional shift in all group specific odds ratio, $D$ would always decline as a result of such balanced increase in opportunity. Hence, out of the three
concepts of balanced increase in opportunity we consider in these notes, $D$ is insensitive to one and inverse related to the other two. One may conclude therefore that $D$ is an inequality measure exhibiting some pro-growth bias.
Table 2: Sensitivity of Alternative measures of inequality of opportunity to a balanced increase in available opportunities

<table>
<thead>
<tr>
<th>Concept of balanced increase in opportunity</th>
<th>Inequality Measure</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>D</td>
</tr>
<tr>
<td>( p_j^* = (1 + \lambda)p_j )</td>
<td>remains unchanged</td>
</tr>
<tr>
<td>( p_j^* = p_j + \lambda )</td>
<td>decreases</td>
</tr>
<tr>
<td>( p_j^* = \frac{1 + \lambda}{1 + \lambda p_j} p_j )</td>
<td>decreases</td>
</tr>
</tbody>
</table>

Decomposability of the inequality of opportunity index

The degree of inequality of opportunity in this report is measured by

\[
D = \frac{1}{2\bar{p}} \sum_{i=1}^{m} \beta_i |p_i - \bar{p}|
\]

If we denote by \( \nu_i \) the group \( i \) absolute gap in specific coverage relative to the overall mean, then \( \nu_i = |p_i - \bar{p}| \) and,

\[
D = \frac{1}{2\bar{p}} \sum_{i=1}^{m} \beta_i \nu_i
\]

This expression for \( D \) can be extremely useful to disentangle the immediate reasons for variations over time or across countries in the inequality of opportunity. It reveals that \( D \) has only three immediate determinants: \((\bar{p}, \beta, \nu)\), where \( \beta = (\beta_1, \ldots, \beta_m) \) and similarly \( \nu = (\nu_1, \ldots, \nu_m) \). As this expression indicates, \( D \) can change if and only if at least one of these tree determinants changes. For this reason we refer to them as the immediate determinants of the inequality of opportunity.

Usually we associate inequality of opportunity with coverage gaps between circumstance groups. The component of \( D \) capturing these gaps is \( \nu \). The greater the gaps, the greater will tend to be the elements of \( \nu \) and, as consequence, the greater will be \( D \), the inequality of opportunity. The effect of changes in \( \nu \) on changes on \( D \) will be called the “gap effect”.

The factor \( \bar{p} \) appears in the expression to translate an absolute index of inequality into a relative index. It is essential to isolate changes in scale. Without the division by \( \bar{p} \), a balanced increase in the available opportunities (i.e., and increase in which the new
opportunities were distributed precisely as previously available) would increase the level of inequality of opportunity. Moreover, without the division by $\bar{p}$, an increase in available opportunities proportionally assigned would not reduce inequality. It would actually leave inequality unchanged. The effect of changes in $\bar{p}$ on changes on $D$ will be called the “scale effect”.

Finally, the component $\beta$ captures the distribution of individuals across circumstance groups. In the making of $D$, it is the component that allows a vector of coverage gaps $\nu$ to be translated in a scalar measure of inequality. In this construction, $\beta$ serves as the weights of a weighted average. Precisely for this reason, variations in $D$ can be caused to some extent by changes in these weights. It may be useful to isolate them because they refer to changes in the relative size of each circumstance group, not to changes in gaps between groups nor to changes in available opportunities. Hence, the effect of changes in $\beta$ on changes on $D$ will be called the “composition effect”.

Consider two situations: $A$ and $B$. It could be a country at two points in time or two countries at the same point in time. Let $\Delta_{\bar{p}}$ be the changes on $D$ if the only difference between situations $A$ and $B$ were a difference in their $\bar{p}$ s. The $\beta$s and $\nu$s are exactly the same. So $\Delta_{\bar{p}}$ are the changes on $D$ if the only difference between situations $A$ and $B$ were a difference in their $\beta$s, and $\Delta_{\nu}$ the changes on $D$ if the only difference were a difference in their $\nu$s.

Next, we consider a procedure to decompose variations in the inequality of opportunity to concomitant variations in each of its three immediate determinants. Let $\Delta_{\bar{p}}$, $\Delta_{\beta}$ and $\Delta_{\nu}$ be defined as

$$\Delta_{\bar{p}} = \frac{1}{2\bar{p}^A} \sum_{j=1}^{m} \beta_j^A \nu_j^B - \frac{1}{2\bar{p}^B} \sum_{j=1}^{m} \beta_j^B \nu_j^B$$

$$\Delta_{\beta} = \frac{1}{2\bar{p}^A} \sum_{j=1}^{m} \beta_j^A \nu_j^B - \frac{1}{2\bar{p}^A} \sum_{j=1}^{m} \beta_j^A \nu_j^B$$

$$\Delta_{\nu} = \frac{1}{2\bar{p}^A} \sum_{j=1}^{m} \beta_j^A \nu_j^B - \frac{1}{2\bar{p}^B} \sum_{j=1}^{m} \beta_j^B \nu_j^B$$

Notice that these differentials add to the total change on the degree of inequality of opportunity, $\Delta = D^A - D^B$. In fact,

$$\Delta_{\bar{p}} + \Delta_{\beta} + \Delta_{\nu} = \frac{1}{2\bar{p}^A} \sum_{j=1}^{m} \beta_j^A \nu_j^B - \frac{1}{2\bar{p}^A} \sum_{j=1}^{m} \beta_j^A \nu_j^B = D^A - D^B = \Delta$$
so, \( \Delta_p + \Delta_\beta + \Delta_\nu = \Delta \). Moreover, notice that (a) if \( \bar{p}' = \bar{p}'^{B} \) then \( \Delta_p = 0 \), (b) if \( \alpha'^A = \alpha'^B \) then \( \Delta_\alpha = 0 \) and finally, (c) if \( \nu'^A = \nu'^B \) then \( \Delta_\nu = 0 \).

Properties (a)-(c) indicates that each differential is in fact capturing only the impact of the corresponding immediate determinant. In addition, \( \Delta_p + \Delta_\beta + \Delta_\nu = \Delta \). These differentials add to the total change on the degree of inequality of opportunity, \( \Delta = D'^A - D'^B \). Hence the impact of the three immediate determinants adds up to the total change, leading, therefore, to a proper decomposition of the overall change.

3. FROM COVERAGE RATE AND INEQUALITY OF OPPORTUNITIES TO AN OPPORTUNITY INDEX

Everyone recognizes that for equality of opportunity to prevail and poverty to be eradicated, access to basic opportunities should be universal. Let \( I \) be an indicator of access to a basic opportunity (access to primary education of quality, for instance), so that \( I=1 \) for those with access and \( I=0 \) for those without access to this opportunity. Since equality of opportunity would require access for all, a natural indicator to monitor is the coverage rate (i.e., the percentage of subjects with access to this opportunity), \( \bar{p} = \frac{P[I=1]}{E[I]} \).

Figure 1 presents estimates of the current access to primary education of quality in a variety of Latin American and Caribbean countries. We measure the access to this opportunity by the completion of sixth grade on time. According to Figure 3, the access to primary education of quality varies from less than 50% in Guatemala and Nicaragua to near 90% in Jamaica and Chile.
Figure 3: Average probability of completing the sixth grade on time – National, circa 2005

Despite the central importance of this indicator (coverage rate), its total lack of sensitivity to how available opportunities are distributed may be considered an important drawback. It could be argued that the way the available opportunities are distributed should be reflected in any meaningful index of opportunity. Countries with the same coverage would have different opportunities depending on whether the available opportunities are distributed according race, other family origin and socio-economic conditions (circumstances) instead of randomly.

Probably most of the interest and motivation for incidence analysis comes from concerns of this nature. Figure 4 presents how our measure of access to primary education of quality varies for children in poor and rich families. This figure reveals that in addition of lower coverage, in Nicaragua and Guatemala the gap in access between the rich and poor is much greater than in Chile and Jamaica.
If $z$ is a scalar indicator of one of the dimensions of personal origin, gender or family socio-economic background (per capita income in Figure 2), traditional incidence analysis seek to estimate $p(z) = P[I = 1|z] = E[D|z]$ and evaluate how sensitive to $z$ is $p(z)$. Although no general measure of the dependence of $p(z)$ on $z$ is proposed, the ratio between the specific coverage rate for the top and bottom quintiles are widely used. Figure 5 presents estimates for this ratio.
To the extent these gaps in access due to differences in race, gender, family origin and socio-economic condition are important, they should also be part of a general measure of the availability of opportunities. What we need then is a distributive sensitive coverage indicator. Before we attempt to construct such an index, let us solve first certain limitations of the most basic incidence analysis.

A sizeable fraction of incidence analysis studies investigate the relation between access and each circumstance separately, without considering the impact of others circumstances or even holding them constant. These studies seek correlations not causal relations. Figures 4 and 5 certainly demonstrate that access to primary education of quality remains quite correlated with income in almost all countries of the region. The statistical analysis behind this table, however, cannot identify whether the gaps are really caused by differences in income or to other associated variables, like parents education.

One alternative is to estimate multivariate models relating the probability of access to all measurable circumstances. One hopes that proceeding this way would lead closer to identifying the actual causal impact of each circumstance on the probability of access. Needless to say that due to either (a) our limited ability of measuring all circumstances, (b) the existence of other factors, beyond circumstances, affecting the probability of access or (c) the measurement error associated to circumstances the estimated relation between the probability of access and observed circumstances may end up being quite apart from their true causal relation. Nevertheless, expecting to get closer to a causal
relationship, we proceed to a multivariate estimation. More specifically, we estimate 
\[ p(x) = P[I = 1|x] = E[I|x] \] where now \( x \) denotes a comprehensive vector of circumstances.

Based on this estimated relation a large variety of differentials could be computed. Figure 6 and 7 illustrate this possibility. In this figure we contrast the access to primary education of quality for children in families with favorable and unfavorable circumstances. In this comparison a variety of circumstances are varying at the same time. It may be worthwhile to isolate the impact of each dimension of the circumstances, holding all others constant. This can be easily done. Figures 8a to 8c illustrate it.

**Figure 6: Differential probability of completing sixth grade on time, circa 2005**
There is not, however, a unique way for doing it, since the relationships are non-linear and the other circumstances could be held constant at different levels. In Figures 6a-c we opt for taking the average of all these possibilities. More precisely, we estimate

\[
q(x_k) = \int p(x) dF(x_{-k}),
\]

where \(k\) denotes the dimension of immediate interest and \(-k\) all the others\(^{12}\). If the dimension of interest were gender, then all we have to do would be to estimate for every child two access probabilities. In one case assuming that he or she were a male; in the other assuming that he or she were a female. The average of the first probability would be \(q\) for males and the average of the second probability would be \(q\) for females.

\[^{12}\text{Note that } p(x_k) = E[D|x_k], \text{ the univariate relation between the probability of access and the dimension } k \text{ of the circumstance, is given by } p(x_k) = \int p(x) dF(x_{-k}|x_k). \text{ Therefore, in general } q(x_k) \neq p(x_k), \text{ unless } F(x_{-k}|x_k) = F(x_{-k}), \text{ i.e., circumstance } x_k \text{ is stochastically independent of all others.}\]


Figure 8.a: Probability of completing the sixth grade on time by parents education

Figure 8.b: Probability of completing the sixth grade on time by per capita family income
Figure 8.c: Probability of completing the sixth grade on time accordingly the number of siblings

The movement to a multidimensional relation has hopefully led us closer to causal relationship, but it has a major practical drawback. We now have a very large number of probability gaps related to differences in circumstances. How should them be summarized? And once having found a summary statistics, how should we use this statistics to construct a distributive sensitive coverage measure?

Let us consider the first problem first. Surely, there are several alternatives to summarize the strength of the relationship between circumstances and the probability of access. Few procedures, however, share a simple and clear interpretation. Among these few, we consider the minimum proportion of all available opportunities that one must reallocated to ensure equal access to all circumstance groups, i.e., a situation in which \( p(x) \) would be equal to \( \overline{p} \).

To construct this index consider a population with \( N \) subjects divided among \( m \) disjoint circumstance groups. Within each of these groups all subjects have the same value of the circumstances, i.e, within group \( j \) all subjects have \( x = a_j \). Moreover, if a subject has \( x = a_j \) then he/she should belong to circumstance group \( j \). Let \( M_j \) denote the number of subjects in group \( j \) with access to the opportunity in study (i.e., for whom \( I = 1 \)) and let \( N_j \) be the total number of subjects in this group. Then, \( p(a_j) = p_j = \frac{M_j}{N_j} \) and let \( \beta_j = \frac{N_j}{N} \) denote the proportion of subjects in circumstance group \( j \). Finally, let \( \overline{M}_j \) denote the number of subjects that would have access to the opportunity in study if the probability of access were identical for all groups. From its very definition, it follows that \( \overline{M}_j = \overline{p}.N_j \)
Relative to a situation in which the access is identical for all circumstance groups, the number of opportunities in excess or lacking in group \( j \) is given by \( |M_j - \overline{M}_j| \). Therefore, the minimum number of opportunities that need to be rearranged to reach equal access across groups is given by \( \frac{1}{2} \sum_{j=1}^{m} |M_j - \overline{M}_j| \). Since the total number of opportunities available is \( M = \sum_{j=1}^{m} M_j = \overline{p}N \), it follows that as a percentage of the total number of available opportunities the minimum amount that need to be rearranged to ensure equal access across circumstance groups, \( D \), is given by \( D = \frac{1}{2\overline{p}N} \sum_{j=1}^{m} |M_j - \overline{M}_j| \).

Notice that this is exactly the same Dissimilatity Index developed in section 2.1 above. This index can be expressed in many different forms. In fact, notice:

\[
D = \frac{1}{2\overline{p}N} \sum_{j=1}^{m} |M_j - \overline{M}_j| = \frac{1}{2 \overline{p}} \sum_{j=1}^{m} \frac{N_j}{N} |\frac{M_j - \overline{M}_j}{N_j}| = \frac{1}{2 \overline{p}} \sum_{j=1}^{m} \beta_j |p_j - \overline{p}|
\]

As this expression indicates, the index is proportion to the average absolute distance of the group specific access probabilities and the overall probability of access. In this sense, it is a measure of the inequality of opportunity.

There is yet another alternative expression for the same index from which a new interpretation is worthwhile to register. Let \( \gamma_j \) denote the proportion of the overall available opportunities that are being currently allocated to group \( j \). Hence, \( \gamma_j = \frac{M_j}{M} \) and we can rewrite \( D \) as

\[
D = \frac{1}{2\overline{p}N} \sum_{j=1}^{m} |M_j - \overline{M}_j| = \frac{1}{2M} \sum_{j=1}^{m} |M_j - \overline{p}N_j| = \frac{1}{2} \sum_{j=1}^{m} \gamma_j - \beta_j
\]

So, \( D \) has the alternative interpretation of being the distance between the distribution of the opportunities, \( \{\gamma_j\} \), and the distribution of the population, \( \{\beta_j\} \), across circumstance groups. Since ideally each member of the population would have an opportunity, the distribution of the population across groups is identical to the distribution that opportunities would have if equality of opportunity prevails. Hence, the proposed measured of inequality of opportunity is identical to the distance between the actual distribution of opportunities across circumstance groups and the distribution that would occur if equality of opportunity prevails. In this sense is a clear measure of the inequality of opportunity.
Figure 9 presents estimates of the inequality of access to primary education of quality for a variety of Latin American and Caribbean countries, using this inequality measure. Consistently with our previous findings, the inequality of opportunity is particularly large in Nicaragua and Guatemala and particularly small in Jamaica and Chile. Quantitatively, this figure reveals that almost 30% of the educational opportunities in Guatemala and Nicaragua would need to be rearranged for equality of opportunity to prevail. In Chile, Jamaica and Argentina only around 3% of the opportunities would need to be rearranged to reach equality of opportunity.

The estimates presented in this figure are just a scalar synthetic measure of the gaps in access to primary education of quality presented in Figures 2 to 6. It is certainly just one possible measure. However, as we try to demonstrate it is a measure with a simple and clear interpretation that can be computed in a very transparent way.

Having proposed a solution to our first problem let us consider the second: How to construct a distributive sensitive opportunity index related to this measure of inequality of opportunity.

Let us begin analyzing further the meaning of $\overline{p}$. Since $\overline{p} = \frac{M}{N}$ and $N$ is the number of opportunities needed to ensure access for all, we can reinterpret $\overline{p}$ as the percentage of the total number of opportunities required for universal access that are actually available. This form of interpreting $\overline{p}$ clarifies both its strengths and its weaknesses. It demonstrates that $\overline{p}$ is certainly a measure of the stock of available opportunities, but it is completed insensitive to how these opportunities are allocated.
These observations give us clear directions for improving upon $\overline{p}$. A simple and intuitive idea would be to modify the numerator so that are considered as valid only those allocated accordingly to the principle of equal opportunity for all. Hence, if we let $O$ denote the available opportunities allocated respecting the principle of equal opportunity, then the desired index, $r$, can be expressed as $r = \frac{O}{N}$.

It remains, however, to be more specific about $O$. One alternative we already have. Since our measure of inequality of opportunity, $D$, is the proportion of opportunities that must be reallocated for equality of opportunity to prevail, then $1-D$ is the proportion properly allocated and $M(1-D)$ the total number allocated accordingly the principle of equal opportunity for all. Hence, one possibility is to let $O = M(1-D)$. In this case, the overall measure of opportunity, $r$, will be given by

$$ r = \frac{O}{N} = \frac{M(1-D)}{N} = \overline{p}(1-D) $$

Using this measure, Figure 10 presents estimates of the access to primary education of quality for a variety of Latin American and Caribbean countries. This figure reveals that in Chile and Jamaica around 85% of the educational opportunities required for universal access are available and have been allocated consistently with the principle of equal opportunity. On the other hand, in Guatemala and Nicaragua less than 30% of the educational opportunities required for universal access are available and have been allocated consistently with the principle of equal opportunity.\(^\text{13}\)

\(^\text{13}\) The Opportunity index for the 19 selected Latin American countries are shown in Table A.3.8 at the end of this paper.
Although a clear definition must be an essential attribute of any opportunity index, its usefulness will also necessarily depend on its properties. Next, we highlight a few important and useful properties of the proposed index.

### 3.2. Properties of the Opportunity Index

**Sensitivity to redistribution**

We begin rewriting the index. In order to simplify the exposition and without any loss of generality we assume that all circumstance groups are ordered increasingly by their probability of access. Hence,

\[
r = \bar{p}(1 - D) = \bar{p} - \frac{1}{2} \sum_{j=1}^{m} \beta_j |p_j - \bar{p}| = \bar{p} - \frac{1}{2} \left( \sum_{j=1}^{l} \beta_j (\bar{p} - p_j) + \sum_{j=l+1}^{m} \beta_j (p_j - \bar{p}) \right)
\]

where \(l\) denote the number of groups with access probability below average. Let \(\lambda\) denote the proportion of subjects in circumstance groups with access probability below average, then \(\lambda = \sum_{j=1}^{l} \beta_j\). Notice that if \(\bar{p}\), in addition to be the average of the group specific probability of access, were also the median, then \(\lambda = 0.5\). In any case,

\[
r = (2 - \lambda) \sum_{j=1}^{l} \beta_j p_j + (1 - \lambda) \sum_{j=l+1}^{m} \beta_j p_j
\]
This expression for the Opportunity Index $r$ reveals several of its properties. First, it demonstrates that despite its distributive sensitivity the index is Pareto consistent. In fact, an increase in the number of opportunities available to any group would always increase the index. Its distributive sensitiveness, however, implies that the impact would be greater if the increase benefits the groups with below average access to opportunities.

In fact, if an extra opportunity is given to any group with access probability above average, the index would increase by $100 \frac{1-\lambda}{N}$ percentage points. Notice that if $\lambda = 0.5$ and $N = 100$, then a one percentage point increase in the availability of opportunities to above average groups would increase the Opportunity Index in one-half of a percentage point.

However if the additional opportunities were given to any group with access probability below average, the index would increase by $100 \frac{2-\lambda}{N}$ percentage points. Notice, again, that if $\lambda = 0.5$ and $N = 100$, then an one percentage point increase in the availability of opportunity to above average groups would increase in the Opportunity Index in 1.5 percentage point.

Finally, it follows from these observations that a balanced increase in available opportunities that give $\lambda$ to below average groups and $1-\lambda$ for the above average groups would increase the Opportunity Index precisely by $100 \frac{1}{N}$ percentage points. Hence if $\lambda = 0.5$ and $N = 100$, then a balanced increase of one-half percentage point in the availability of opportunity to above average groups couple with the same increase for the bellow average group would increase the Opportunity Index by precisely one percentage point.

*Decomposability of changes in the opportunity index*

Consider two situations: $A$ and $B$. It could be a country at two points in time or two countries at the same point in time. Any change in the opportunity index, $O$, could be decomposed in a scale, $\Delta_p$, and a distributional, $\Delta_D$, effect via

$$
\Delta = O^B - O^A = \bar{p}^B (1-D^B) - \bar{p}^A (1-D^A) = \Delta_p + \Delta_D
$$

where the scale effect, $\Delta_p$, and the distributional effect, $\Delta_D$, are defined as follows.
\[ \Delta_p = \bar{p}^a (1 - D^a) - \bar{p}^d (1 - D^d) \]

and

\[ \Delta_D = \bar{p}^a (1 - D^a) - \bar{p}^d (1 - D^d) \]

Moreover, notice that (a) if \( \bar{p}^d = \bar{p}^a \) then \( \Delta_p = 0 \) and (b) if \( D^a = D^d \) then \( \Delta_D = 0 \).

These properties indicate that each differential is in fact capturing only the impact of the corresponding factor. Since, by construction, \( \Delta_p + \Delta_D = \Delta \), the impact of the two determinants of the opportunity index adds up to the total change, leading, therefore, to a proper decomposition.\(^{14}\)

4. SUMMARY

This chapter has provided a broad discussion of the concept of inequality of opportunities and the issues involved with measuring this concept. Measuring inequality of opportunities requires special attention to understanding how circumstances, efforts, innate talents, luck and other factors affect outcomes and how these elements can be appropriately gauged and controlled for given the usually available data. Several approaches to measuring inequality of opportunities are explained, and its relative advantages are discussed in terms of both its relevance and applicability. After these considerations, an index for measuring inequality of opportunities (the Dissimilarity Index, D) is proposed. In addition, an Opportunity Index is proposed that takes into consideration expansion in the coverage of basic opportunities as well as its distribution. Both indexes have a series of properties, fully elaborated in the text, that make them useful for analytical and policymaking purposes. Finally, the indexes are illustrated for a sample of 19 Latin American countries with available data for recent years with respect to access to selected basic opportunities (i.e., primary education, water and sanitation, and electricity).

---

\(^{14}\) Other decomposition formulas are possible.
REFERENCES

### ANNEX 3.1: REFERENCE TABLES

#### Table A.3.1: Logistic regression estimates

| Inequality of opportunity: Probability of completing the sixth grade - circa 1995 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
|   | Bolivia | Brazil | Chile | Colombia | Costa Rica | Dominican Republic | Ecuador | El Salvador | Guatemala | Honduras | Jamaica | Mexico | Nicaragua | Panama | Paraguay |   |
| Age |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| 8 years |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| 9 years |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| 10 years |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| 11 years |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| 12 years |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| 13 years | 1.23 | 2.30 | 2.15 | 1.22 | 2.02 | 0.85 | 1.42 | 1.52 | 1.51 | 1.67 | 2.51 | 1.08 | 0.95 | 1.45 | 1.71 |
| 14 years | 2.16 | 3.19 | 3.09 | 1.79 | 2.70 | 1.68 | 1.86 | 2.35 | 2.32 | 2.67 | 3.92 | 1.52 | 1.59 | 1.99 | 2.51 |
| 15 years | 2.49 | 3.58 | 3.69 | 2.28 | 3.34 | 2.07 | 2.31 | 2.76 | 2.87 | 2.97 | 2.88 | 2.14 | 2.12 | 2.52 | 3.41 |
| 16 years | 2.82 | 3.92 | 4.08 | 2.53 | 3.18 | 2.23 | 2.42 | 3.05 | 3.14 | 3.27 | 1.95 | 1.93 | 2.36 | 2.94 | 3.61 |
| Gender (male) | 0.21 | -0.55 | -0.36 | -0.40 | -0.28 | -0.50 | -0.11 | -0.36 | 0.22 | -0.25 | 0.71 | -0.22 | -0.43 | -0.24 | -0.1 |
| Years of schooling of the family head | 0.16 | 0.28 | 0.17 | 0.26 | 0.26 | 0.14 | 0.32 | 0.11 | 0.31 | 0.20 | 0.11 | 0.23 | 0.12 | 0.22 | 0.2 |
| Years of schooling of the family head to the square | 0.00 | -0.01 | 0.00 | -0.01 | -0.01 | 0.00 | -0.01 | 0.00 | 0.00 | 0.00 | -0.01 | 0.00 | -0.01 | -0.01 | -0.0 |
| Logarithm of the per capita income | 0.27 | 0.64 | 0.29 | 0.15 | 0.15 | 0.26 | 0.13 | 0.28 | 0.37 | 0.27 | 0.07 | 0.32 | 0.28 | 0.31 | 0.21 |
| Presence of family head and spouse | -0.26 | 0.35 | 0.14 | 0.13 | 0.16 | 0.01 | 0.41 | -0.09 | -0.15 | 0.12 | 0.28 | 0.29 | -0.16 | 0.20 | -0.01 |
| Number of children age 0 to 16 in the residence | -0.11 | -0.15 | -0.15 | -0.23 | -0.14 | -0.23 | -0.15 | -0.11 | -0.10 | -0.08 | 0.16 | -0.13 | -0.13 | -0.14 | -0.1 |
| Area (urban) | 0.64 | 0.48 | 0.09 | 0.83 | 0.10 | 0.28 | 0.37 | 0.76 | 0.93 | 0.41 | -0.54 | 0.51 | 1.26 | 0.39 | -0.0 |

**Note:** The table includes logistic regression estimates for various countries and variables related to inequality of opportunity, specifically the probability of completing the sixth grade around 1995. The table lists coefficients for different variables including age, gender, years of schooling, logarithm of per capita income, and presence of family head and spouse.
### Table A.3.2: Logistic regression estimates

| Inequality of opportunity: Probability of completing the sixth grade - circa 2005 | Argentina | Bolivia | Brazil | Chile | Colombia | Costa Rica | Dominican Republic | Ecuador | El Salvador | Guatemala | Honduras | Jamaica | Mexico | Nicaragua | Panama | Paraguay | Peru | Uruguay |
| Intercept | 0.19 | -2.75 | -5.71 | -2.54 | -4.75 | -4.71 | -3.94 | -0.58 | -3.76 | -4.76 | -3.76 | -0.23 | -2.12 | -3.80 | -1.25 | -4.21 | -4.10 | -4.60 |
| Age | | | | | | | | | | | | | | | | | | | | |
| 8 years | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| 9 years | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| 10 years | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| 11 years | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| 12 years | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| 13 years | 0.65 | 1.32 | 2.39 | 2.14 | 0.94 | 2.42 | 1.02 | 0.77 | 1.47 | 2.00 | 1.56 | 2.14 | 1.19 | 1.10 | 1.70 | 1.50 | 1.87 | 1.64 |
| 14 years | -0.33 | 1.87 | 3.28 | 3.35 | 1.52 | 3.37 | 1.54 | 1.15 | 2.53 | 2.81 | 2.45 | 2.89 | 2.00 | 1.70 | 2.31 | 2.35 | 2.85 | 2.73 |
| 15 years | 0.71 | 2.23 | 3.74 | 4.01 | 1.58 | 4.07 | 2.03 | 1.30 | 2.72 | 3.36 | 2.77 | 2.65 | 2.13 | 2.11 | 2.90 | 2.75 | 3.21 | 3.26 |
| 16 years | 1.81 | 2.09 | 4.06 | 4.50 | 1.80 | 3.95 | 2.28 | 1.53 | 3.14 | 3.53 | 2.87 | 2.61 | 2.20 | 2.47 | 2.76 | 2.97 | 3.29 | 3.51 |
| Gender (male) | 0.11 | 0.01 | -0.65 | -0.24 | -0.52 | -0.42 | -0.77 | -0.15 | -0.17 | 0.13 | -0.32 | -0.44 | 0.01 | -0.54 | -0.48 | -0.41 | -0.03 | -0.55 |
| Years of schooling of the family head | 0.17 | 0.25 | 0.21 | 0.15 | 0.17 | 0.24 | 0.18 | 0.22 | 0.17 | 0.26 | 0.20 | -0.02 | 0.22 | 0.18 | 0.17 | 0.23 | 0.28 | 0.22 |
| Years of schooling of the family head to the square | -0.01 | -0.01 | -0.01 | 0.00 | 0.00 | -0.01 | 0.00 | 0.00 | -0.01 | 0.00 | 0.00 | -0.01 | 0.00 | 0.00 | -0.01 | -0.01 | -0.01 | -0.01 |
| Logarithm of the per capita income | -0.01 | 0.23 | 0.46 | 0.13 | 0.36 | 0.24 | 0.36 | 0.25 | 0.24 | 0.28 | 0.10 | 0.33 | 0.35 | 0.16 | 0.25 | 0.29 | 0.48 |
| Presence of family head and spouse | 0.08 | 0.06 | 0.29 | 0.17 | 0.02 | 0.29 | 0.16 | -0.08 | 0.00 | -0.08 | 0.09 | 0.13 | 0.12 | 0.14 | 0.01 | 0.19 | -0.07 | 0.29 |
| Number of children age 0 to 16 in the residence | -0.09 | -0.11 | -0.13 | -0.12 | -0.22 | -0.22 | -0.13 | -0.17 | -0.10 | -0.12 | -0.14 | -0.09 | -0.15 | -0.21 | -0.20 | -0.19 | -0.18 | -0.20 |
| Area (urban) | -0.17 | 0.69 | 0.41 | -0.09 | 0.80 | 0.11 | 0.10 | 0.19 | 0.45 | 0.52 | 0.28 | 0.31 | -0.34 | 0.80 | 0.26 | -0.11 | 0.55 | -0.11 |
### Table A.3.3: Logistic regression estimates

**Inequality of opportunity: Access to electricity - circa 1995**

<table>
<thead>
<tr>
<th></th>
<th>Argentina</th>
<th>Bolivia</th>
<th>Brazil</th>
<th>Chile</th>
<th>Colombia</th>
<th>Costa Rica</th>
<th>Dominican Republic</th>
<th>Ecuador</th>
<th>El Salvador</th>
<th>Guatemala</th>
<th>Honduras</th>
<th>Jamaica</th>
<th>Mexico</th>
<th>Nicaragua</th>
<th>Panama</th>
<th>Paraguay</th>
<th>Peru</th>
<th>Uruguay</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-3.53</td>
<td>-2.52</td>
<td>-1.01</td>
<td>-4.73</td>
<td>-1.24</td>
<td>-4.47</td>
<td>-5.62</td>
<td>0.56</td>
<td>-2.86</td>
<td>-3.02</td>
<td>-4.95</td>
<td>1.67</td>
<td>-0.68</td>
<td>-4.07</td>
<td>-2.28</td>
<td>-3.29</td>
<td>-3.64</td>
<td>-8.68</td>
</tr>
<tr>
<td>Gender (male)</td>
<td>0.36</td>
<td>-0.79</td>
<td>-0.22</td>
<td>-1.01</td>
<td>-0.64</td>
<td>-0.75</td>
<td>-1.16</td>
<td>-0.83</td>
<td>-0.09</td>
<td>-0.40</td>
<td>-0.54</td>
<td>0.41</td>
<td>-0.92</td>
<td>-0.58</td>
<td>0.11</td>
<td>-0.17</td>
<td>-0.23</td>
<td></td>
</tr>
<tr>
<td>Years of schooling of the family head</td>
<td>-0.16</td>
<td>-0.14</td>
<td>-0.08</td>
<td>0.10</td>
<td>0.13</td>
<td>0.17</td>
<td>0.09</td>
<td>0.12</td>
<td>0.13</td>
<td>0.05</td>
<td>-0.13</td>
<td>0.23</td>
<td>0.06</td>
<td>0.09</td>
<td>0.12</td>
<td>0.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Logarithm of the per capita income</td>
<td>0.32</td>
<td>0.33</td>
<td>0.50</td>
<td>0.64</td>
<td>0.30</td>
<td>0.76</td>
<td>0.87</td>
<td>0.48</td>
<td>0.60</td>
<td>0.67</td>
<td>0.74</td>
<td>0.07</td>
<td>0.65</td>
<td>0.62</td>
<td>0.84</td>
<td>0.90</td>
<td>0.61</td>
<td>1.27</td>
</tr>
<tr>
<td>Presence of family head and spouse</td>
<td>0.24</td>
<td>0.22</td>
<td>0.20</td>
<td>1.10</td>
<td>0.12</td>
<td>0.30</td>
<td>0.78</td>
<td>0.36</td>
<td>0.08</td>
<td>-0.19</td>
<td>0.19</td>
<td>0.01</td>
<td>0.97</td>
<td>-0.07</td>
<td>-0.18</td>
<td>0.15</td>
<td>-0.02</td>
<td>0.50</td>
</tr>
<tr>
<td>Number of children age 0 to 16 in the residence</td>
<td>0.55</td>
<td>-0.20</td>
<td>-0.10</td>
<td>-0.06</td>
<td>0.02</td>
<td>-0.09</td>
<td>0.19</td>
<td>-0.14</td>
<td>-0.31</td>
<td>-0.02</td>
<td>0.00</td>
<td>-0.12</td>
<td>-0.21</td>
<td>-0.05</td>
<td>-0.02</td>
<td>-0.13</td>
<td>-0.02</td>
<td></td>
</tr>
<tr>
<td>Area (urban)</td>
<td>1.34</td>
<td>2.44</td>
<td>3.07</td>
<td>2.63</td>
<td>0.10</td>
<td>4.46</td>
<td>1.06</td>
<td>2.91</td>
<td>1.42</td>
<td>1.06</td>
<td>2.92</td>
<td>0.78</td>
<td>0.96</td>
<td>2.77</td>
<td>2.16</td>
<td>1.30</td>
<td>2.88</td>
<td>2.79</td>
</tr>
</tbody>
</table>

### Table A.3.4: Logistic regression estimates

**Inequality of opportunity: Access to electricity - circa 2005**

<table>
<thead>
<tr>
<th></th>
<th>Argentina</th>
<th>Bolivia</th>
<th>Brazil</th>
<th>Chile</th>
<th>Colombia</th>
<th>Costa Rica</th>
<th>Dominican Republic</th>
<th>Ecuador</th>
<th>El Salvador</th>
<th>Guatemala</th>
<th>Honduras</th>
<th>Jamaica</th>
<th>Mexico</th>
<th>Nicaragua</th>
<th>Panama</th>
<th>Paraguay</th>
<th>Peru</th>
<th>Uruguay</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.80</td>
<td>-2.92</td>
<td>-2.92</td>
<td>-4.73</td>
<td>-1.24</td>
<td>-4.47</td>
<td>-5.62</td>
<td>0.56</td>
<td>-2.86</td>
<td>-3.02</td>
<td>-4.95</td>
<td>1.67</td>
<td>-0.68</td>
<td>-4.07</td>
<td>-2.28</td>
<td>-3.29</td>
<td>-3.64</td>
<td>-8.68</td>
</tr>
<tr>
<td>Gender (male)</td>
<td>0.36</td>
<td>-0.79</td>
<td>-0.22</td>
<td>-1.01</td>
<td>-0.64</td>
<td>-0.75</td>
<td>-1.16</td>
<td>-0.83</td>
<td>-0.09</td>
<td>-0.40</td>
<td>-0.54</td>
<td>0.41</td>
<td>-0.92</td>
<td>-0.58</td>
<td>0.11</td>
<td>-0.17</td>
<td>-0.23</td>
<td></td>
</tr>
<tr>
<td>Years of schooling of the family head</td>
<td>-0.16</td>
<td>-0.14</td>
<td>-0.08</td>
<td>0.10</td>
<td>0.13</td>
<td>0.17</td>
<td>0.09</td>
<td>0.12</td>
<td>0.13</td>
<td>0.05</td>
<td>-0.13</td>
<td>0.23</td>
<td>0.06</td>
<td>0.09</td>
<td>0.12</td>
<td>0.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Logarithm of the per capita income</td>
<td>0.32</td>
<td>0.33</td>
<td>0.50</td>
<td>0.64</td>
<td>0.30</td>
<td>0.76</td>
<td>0.87</td>
<td>0.48</td>
<td>0.60</td>
<td>0.67</td>
<td>0.74</td>
<td>0.07</td>
<td>0.65</td>
<td>0.62</td>
<td>0.84</td>
<td>0.90</td>
<td>0.61</td>
<td>1.27</td>
</tr>
<tr>
<td>Presence of family head and spouse</td>
<td>0.24</td>
<td>0.22</td>
<td>0.20</td>
<td>1.10</td>
<td>0.12</td>
<td>0.30</td>
<td>0.78</td>
<td>0.36</td>
<td>0.08</td>
<td>-0.19</td>
<td>0.19</td>
<td>0.01</td>
<td>0.97</td>
<td>-0.07</td>
<td>-0.18</td>
<td>0.15</td>
<td>-0.02</td>
<td>0.50</td>
</tr>
<tr>
<td>Number of children age 0 to 16 in the residence</td>
<td>0.55</td>
<td>-0.20</td>
<td>-0.10</td>
<td>-0.06</td>
<td>0.02</td>
<td>-0.09</td>
<td>0.19</td>
<td>-0.14</td>
<td>-0.31</td>
<td>-0.02</td>
<td>0.00</td>
<td>-0.12</td>
<td>-0.21</td>
<td>-0.05</td>
<td>-0.02</td>
<td>-0.13</td>
<td>-0.02</td>
<td></td>
</tr>
<tr>
<td>Area (urban)</td>
<td>1.34</td>
<td>2.44</td>
<td>3.07</td>
<td>2.63</td>
<td>0.10</td>
<td>4.46</td>
<td>1.06</td>
<td>2.91</td>
<td>1.42</td>
<td>1.06</td>
<td>2.92</td>
<td>0.78</td>
<td>0.96</td>
<td>2.77</td>
<td>2.16</td>
<td>1.30</td>
<td>2.88</td>
<td>2.79</td>
</tr>
</tbody>
</table>
### Table A.3.5: Logistic regression estimates

| Inequality of opportunity: Adequate access to water and sanitation - circa 1995 | Argentina | Bolivia | Brazil | Chile | Colombia | Costa Rica | Dominica Republic | Ecuador | El Salvador | Guatemala | Honduras | Jamaica | Mexico | Nicaragua | Panama | Paraguay | Peru | Uruguay | Venezuela |
| Intercept | -5.37 | -5.88 | -12.92 | -3.40 | -5.47 | -9.21 | -5.79 | -4.50 | -1.27 | -6.95 | -8.45 | -4.01 | -13.56 | -8.03 | -4.90 |
| Gender (male) | -0.12 | -0.22 | -0.39 | -0.38 | -0.21 | -0.66 | -0.07 | 0.39 | -0.40 | -0.32 | 0.19 | -0.40 | 0.48 | -0.38 |
| Years of schooling of the family head | 0.02 | 0.06 | 0.18 | 0.07 | 0.31 | 0.05 | 0.02 | 0.03 | 0.15 | 0.00 | -0.10 |
| Years of schooling of the family head to the square | 0.00 | 0.00 | 0.00 | 0.03 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.00 | 0.00 |
| Logarithm of the per capita income | 0.30 | 0.86 | 0.57 | 0.55 | -0.12 | 0.93 | 0.61 | 0.52 | 0.94 | 0.89 | -0.66 |
| Presence of family head and spouse | 0.30 | 0.86 | 0.57 | 0.55 | -0.12 | 0.93 | 0.61 | 0.52 | 0.94 | 0.89 | -0.66 |
| Number of children age 0 to 16 in the residence | -0.12 | 0.01 | -0.10 | 0.01 | -0.04 | -0.05 | 0.03 | -0.05 | -0.06 | 0.03 | -0.03 |
| Area (urban) | 1.48 | 3.44 | 2.49 | 0.89 | 2.28 | 2.02 | 4.27 | 2.06 | 1.80 | 3.28 | -0.70 |

### Table A.3.6: Logistic regression estimates

| Inequality of opportunity: Adequate access to water and sanitation - circa 2005 | Argentina | Bolivia | Brazil | Chile | Colombia | Costa Rica | Dominica Republic | Ecuador | El Salvador | Guatemala | Honduras | Jamaica | Mexico | Nicaragua | Panama | Paraguay | Peru | Uruguay | Venezuela |
| Gender (male) | -0.53 | -0.44 | -0.20 | 0.01 | -0.33 | -0.26 | 0.02 | 0.09 | -0.07 | -0.61 | -0.10 | 0.68 | -0.04 | -0.02 | 0.12 | 0.06 | 0.03 | 0.18 |
| Years of schooling of the family head | 0.08 | 0.04 | 0.20 | -0.12 | 0.11 | 0.15 | 0.09 | 0.09 | 0.05 | 0.14 | 0.18 | -0.13 | 0.07 | 0.21 | 0.00 | 0.19 | -0.05 | 0.11 |
| Years of schooling of the family head to the square | 0.01 | 0.01 | 0.00 | 0.00 | 0.01 | 0.00 | 0.00 | 0.00 | 0.00 | 0.01 | 0.00 | -0.01 | 0.01 | 0.00 | 0.00 | 0.01 | 0.00 |
| Logarithm of the per capita income | 0.54 | 0.44 | 0.61 | 0.88 | 0.47 | 0.94 | 0.88 | 0.52 | 0.74 | 0.78 | 0.52 | 0.90 | 0.74 | 0.71 | 1.03 | 0.72 | 1.43 |
| Presence of family head and spouse | 0.07 | 0.10 | -0.10 | -0.44 | -0.79 | -0.11 | -0.45 | -0.31 | -0.34 | 0.29 | -0.13 | -0.29 | -0.02 | -0.45 | -0.65 | -0.49 | -0.49 | -0.14 |
| Number of children age 0 to 16 in the residence | -0.12 | -0.12 | 0.01 | 0.11 | -0.07 | -0.16 | 0.01 | -0.11 | 0.00 | -0.08 | 0.03 | -0.12 | -0.05 | -0.02 | -0.07 | -0.14 | -0.07 | -0.04 |
| Area (urban) | 0.28 | 2.05 | 2.14 | 3.00 | 2.08 | 1.03 | 1.51 | 1.73 | 2.54 | 2.24 | 1.82 | 1.88 | 1.05 | 3.28 | 1.53 | 1.80 | 2.16 | 4.29 |
Table A.3.7: Index of Inequality of Opportunity (D) in 19 LAC Countries

<table>
<thead>
<tr>
<th>Country</th>
<th>Completion of 6th grade on time</th>
<th>Electricity</th>
<th>Water and sanitation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Argentina</td>
<td>...........</td>
<td>3%</td>
<td>...........</td>
</tr>
<tr>
<td>Bolivia</td>
<td>16%</td>
<td>12%</td>
<td>3%</td>
</tr>
<tr>
<td>Brazil</td>
<td>36%</td>
<td>20%</td>
<td>16%</td>
</tr>
<tr>
<td>Chile</td>
<td>6%</td>
<td>3%</td>
<td>4%</td>
</tr>
<tr>
<td>Colombia</td>
<td>19%</td>
<td>15%</td>
<td>5%</td>
</tr>
<tr>
<td>Costa Rica</td>
<td>12%</td>
<td>9%</td>
<td>3%</td>
</tr>
<tr>
<td>Dominican Republic</td>
<td>16%</td>
<td>13%</td>
<td>3%</td>
</tr>
<tr>
<td>Ecuador</td>
<td>10%</td>
<td>6%</td>
<td>4%</td>
</tr>
<tr>
<td>El Salvador</td>
<td>25%</td>
<td>20%</td>
<td>5%</td>
</tr>
<tr>
<td>Guatemala</td>
<td>37%</td>
<td>27%</td>
<td>10%</td>
</tr>
<tr>
<td>Honduras</td>
<td>22%</td>
<td>19%</td>
<td>3%</td>
</tr>
<tr>
<td>Jamaica</td>
<td>3%</td>
<td>2%</td>
<td>1%</td>
</tr>
<tr>
<td>Mexico</td>
<td>10%</td>
<td>5%</td>
<td>4%</td>
</tr>
<tr>
<td>Nicaragua</td>
<td>34%</td>
<td>24%</td>
<td>10%</td>
</tr>
<tr>
<td>Panama</td>
<td>11%</td>
<td>8%</td>
<td>2%</td>
</tr>
<tr>
<td>Paraguay</td>
<td>15%</td>
<td>11%</td>
<td>4%</td>
</tr>
<tr>
<td>Peru</td>
<td>26%</td>
<td>17%</td>
<td>9%</td>
</tr>
<tr>
<td>Uruguay</td>
<td>...........</td>
<td>7%</td>
<td>...........</td>
</tr>
<tr>
<td>Venezuela</td>
<td>10%</td>
<td>6%</td>
<td>4%</td>
</tr>
<tr>
<td>Average</td>
<td>18%</td>
<td>12%</td>
<td>...........</td>
</tr>
</tbody>
</table>
### Table A.3.8: Opportunity Index in 19 LAC Countries

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Argentina</td>
<td>......</td>
<td>0.73</td>
<td>......</td>
<td>0.75</td>
<td>......</td>
<td>0.75</td>
<td>0.98</td>
<td>0.75</td>
<td>......</td>
</tr>
<tr>
<td>Bolivia</td>
<td>0.50</td>
<td>0.61</td>
<td>0.11</td>
<td>0.23</td>
<td>0.19</td>
<td>-0.04</td>
<td>0.45</td>
<td>0.81</td>
<td>0.92</td>
</tr>
<tr>
<td>Brazil</td>
<td>0.15</td>
<td>0.37</td>
<td>0.22</td>
<td>0.39</td>
<td>0.50</td>
<td>0.11</td>
<td>0.81</td>
<td>0.92</td>
<td>0.05</td>
</tr>
<tr>
<td>China</td>
<td>0.73</td>
<td>0.83</td>
<td>0.09</td>
<td>0.66</td>
<td>0.78</td>
<td>0.12</td>
<td>0.93</td>
<td>0.98</td>
<td>0.05</td>
</tr>
<tr>
<td>Colombia</td>
<td>0.43</td>
<td>0.56</td>
<td>0.13</td>
<td>0.48</td>
<td>0.87</td>
<td>.....</td>
<td>0.86</td>
<td>0.98</td>
<td>.....</td>
</tr>
<tr>
<td>Costa Rica</td>
<td>0.55</td>
<td>0.65</td>
<td>0.10</td>
<td>0.87</td>
<td>0.98</td>
<td>.....</td>
<td>0.89</td>
<td>0.98</td>
<td>.....</td>
</tr>
<tr>
<td>Dominican Republic</td>
<td>0.45</td>
<td>0.57</td>
<td>0.12</td>
<td>0.36</td>
<td>0.89</td>
<td>.....</td>
<td>0.89</td>
<td>0.98</td>
<td>.....</td>
</tr>
<tr>
<td>Ecuador</td>
<td>0.64</td>
<td>0.78</td>
<td>0.13</td>
<td>0.23</td>
<td>0.42</td>
<td>0.19</td>
<td>0.75</td>
<td>0.90</td>
<td>0.15</td>
</tr>
<tr>
<td>El Salvador</td>
<td>0.28</td>
<td>0.34</td>
<td>0.07</td>
<td>0.16</td>
<td>0.24</td>
<td>0.08</td>
<td>0.65</td>
<td>0.75</td>
<td>0.10</td>
</tr>
<tr>
<td>Guatemala</td>
<td>0.16</td>
<td>0.24</td>
<td>0.09</td>
<td>0.10</td>
<td>0.18</td>
<td>0.08</td>
<td>0.57</td>
<td>0.66</td>
<td>0.10</td>
</tr>
<tr>
<td>Honduras</td>
<td>0.28</td>
<td>0.38</td>
<td>0.10</td>
<td>0.22</td>
<td>0.89</td>
<td>.....</td>
<td>0.51</td>
<td>0.89</td>
<td>0.08</td>
</tr>
<tr>
<td>Jamaica</td>
<td>0.87</td>
<td>0.86</td>
<td>-0.01</td>
<td>0.11</td>
<td>0.10</td>
<td>-0.02</td>
<td>0.74</td>
<td>0.83</td>
<td>0.09</td>
</tr>
<tr>
<td>Mexico</td>
<td>0.68</td>
<td>0.80</td>
<td>0.12</td>
<td>0.28</td>
<td>0.46</td>
<td>0.18</td>
<td>0.90</td>
<td>0.97</td>
<td>0.06</td>
</tr>
<tr>
<td>Nicaragua</td>
<td>0.17</td>
<td>0.33</td>
<td>0.16</td>
<td>0.09</td>
<td>0.10</td>
<td>0.02</td>
<td>0.53</td>
<td>0.49</td>
<td>-0.03</td>
</tr>
<tr>
<td>Panama</td>
<td>0.67</td>
<td>0.70</td>
<td>0.03</td>
<td>0.25</td>
<td>0.29</td>
<td>0.04</td>
<td>0.53</td>
<td>0.59</td>
<td>0.06</td>
</tr>
<tr>
<td>Paraguay</td>
<td>0.45</td>
<td>0.59</td>
<td>0.14</td>
<td>0.29</td>
<td>0.38</td>
<td>0.10</td>
<td>0.83</td>
<td>0.91</td>
<td>0.08</td>
</tr>
<tr>
<td>Peru</td>
<td>0.25</td>
<td>0.45</td>
<td>0.19</td>
<td>0.20</td>
<td>0.32</td>
<td>0.12</td>
<td>0.42</td>
<td>0.52</td>
<td>0.10</td>
</tr>
<tr>
<td>Uruguay</td>
<td>......</td>
<td>0.75</td>
<td>......</td>
<td>0.66</td>
<td>......</td>
<td>0.97</td>
<td>0.97</td>
<td>0.97</td>
<td>.....</td>
</tr>
<tr>
<td>Venezuela</td>
<td>0.62</td>
<td>0.73</td>
<td>0.11</td>
<td>0.73</td>
<td>0.77</td>
<td>0.04</td>
<td>0.98</td>
<td>0.98</td>
<td>0.00</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td><strong>0.46</strong></td>
<td><strong>0.59</strong></td>
<td><strong>0.11</strong></td>
<td><strong>0.31</strong></td>
<td><strong>0.43</strong></td>
<td><strong>-0.05</strong></td>
<td><strong>0.70</strong></td>
<td><strong>0.80</strong></td>
<td><strong>-0.05</strong></td>
</tr>
</tbody>
</table>