Ethnicity and Earnings in Urban Peru

Hugo Ñopo
Jaime Saavedra
Maximo Torero

I Z A

Ethnicity and Earnings in Urban Peru

Hugo Ñopo
*Middlebury College, GRADE
and IZA Bonn*

Jaime Saavedra
World Bank, GRADE

Maximo Torero
IFPRI, GRADE

Discussion Paper No. 980
January 2004

ABSTRACT

Ethnicity and Earnings in Urban Peru[*]

In this paper we study the relationship between ethnic exclusion and earnings in Urban Peru. Our approach to the concept of ethnicity involves the usage of instruments in many of its several dimensions: mother tongue, parental background, religion, migration events and race. In order to approximate what can be called racial differences in a context like the Peruvian in which "racial mixture" is the main characteristic of the population, we use a score-based procedure to capture both the differences and the mixtures. By means of this procedure each individual is assigned intensities by pollsters in each of the four categories that correspond to the most easily recognized distinct racial groups in the Peruvian society: Asiatic, White, Indigenous, and Black. We find that the multidimensional race indicator is correlated with several human capital and physical capital assets, as well as with access to public services. Using an extension of the Blinder-Oaxaca (B-O) decompositions to allow comparisons among infinitely many groups, we find that a substantial part of the earnings differences between racial groups cannot be explained by differences in individual characteristics. The results suggests that among wage earners after controlling for a large set of characteristics, there are racially related earnings differences in favor of predominantly White individuals. In the case of the self-employed, none of the empirical distributions of earning differences attributable to race is substantially above zero.

JEL Classification: J15, J31, J71

Keywords: race discrimination, minorities, wage differentials, semi-parametric

Corresponding author:

Hugo R. Ñopo
Warner 305C
Economics Department
Middlebury College
Middlebury, Vermont 05753
USA
Email: hnopo@middlebury.edu

1   Introduction

The outcomes of discrimination and exclusion related to ethnicity, culture, physical appearance and religion are very notorious. At the same time, the mechanisms by which those operate are rather subtle. Indigenous or ethnic minorities are more likely to be poor than any other group. While overall the poverty rate is 54%, according to the 2000 Living Standards Measurement Survey (LSMS), the poverty rate of the population whose mother's tongue is Quechua, Aymara or other native language is 70%. Moreover, more than 75% of this group can be found in the three bottom deciles of the income distribution.

Ethnic and race discrimination in Peru has been studied, usually through the analysis of case studies, rather than in a systematic general approach with some pretension of statistical significance. Callirgos(1993) gives a global overview of the origins and particular characteristics that Peruvian racism may have. Oliart (1989), Pozzi-Escot (1989), Callirgos (1993), and Mendoza (1993) tried to tackle ethnic and cultural discrimination. Finally, using case studies Sulmont (1995) has documented some of the elements of social exclusion that may be present in Peruvian labor markets.

Despite the obvious importance of the topic for a country like Peru, there are very few data sources that can capture ethnic discrimination, and empirical work that tries to tackle exclusion and discrimination issues from a quantitative perspective is scarce. Most of the empirical literature has approximated racial and ethnic discrimination with supposedly easily observable variables, in most of the cases, mother tongue as in several of the World Bank studies. MacIsaac (1993), for instance, finds that more than 80% of non-indigenous people -defined as those whose mother tongue is Spanish- have access to public water supply or access to electricity while less than 45% of indigenous people -defined as those whose mother tongue is Quechua or Aymara- has access to the same type of public goods. She also finds that years of schooling is 8.1 for non-indigenous people while it is only 5.5 for indigenous people, just to mention two of the most striking differences. Psacharopoulos and Patrinos (1994) find that ceteris paribus, individuals whose mother tongue is Quechua have earnings that are 8% lower than the average of the population.

However, the approximation of ethnicity using mother tongue is clearly incomplete, as there are other ethnic differences within the Spanish and Quechua speaking populations.[1] In this paper, in measuring the effect of ethnic exclusion over earnings, we improve over the measuring of discrimination by approximating ethnicity using variables related to several dimensions of the concept, such as mother tongue, parental background, race, and religion. All these variables are linked to differences among individuals that may have measurable consequences over economic opportunities.

---

[1] Other attempts to approximate ethnic characteristics are found in some household surveys in Peru. As an example, the 2000 Living Standards Measurement Survey (LSMS 2000) inquiries about racial characteristics, and more than 98% of the rural populations is self-reported as "Mestiza" (mixed race), and only a tiny percentage of the sample self-report themselves as Indigenous, Asian, Black or White.

One of the major concerns when measuring race is that the Peruvian population is neither predominantly Indigenous nor White nor Black nor Asian, but is a continuum of different degrees of mixtures that is difficult to treat empirically. In order to approximate what can be called racial intensities, a score-based procedure is used. In such procedure every individual receives a score (in an ordinal scale ranging from 0 to 10) for each of the four categories representing the groups that are most easily recognized by people as distinct racial groups: Asian, White, Indigenous, and Black. We construct indicators based on self-reported data and pollster's data.[2] Several papers have looked at the effects of racial differences in a similar fashion. For instance; Arce, Murguia and Frisbie (1987) use in their study two phenotypical dimensions: skin color, ranging from very light to very dark and physical features, from Very European to Very Indian, both on a 1 to 5 scale. Johnson and Farrell (1995) look at differences in income between light-skinned and dark-skinned black males in Los Angeles. Darity and Mason (1998) review studies that use "skin shade", i.e. dark skinned black males with light skinned black males as categories of analysis. Using Latin American data, Silva (1985) compares earnings of blacks, mulatos and whites in Brazil, and Telles and Lim (1998) use a classification of black, morenos and whites to analyze earning differentials in Brazil combining self-reported data with data reported by pollsters.

The instruments that proxy ethnicity are then used in an econometric setup in order to explore their interaction with earnings. For that purpose we extend the Blinder-Oaxaca decomposition of wage gaps in a setup in which there is a continuum of comparing groups, as it is the case for the racial characteristics of the Peruvian population.

The rest of the paper is organized as follows: In Section 2 we introduce our approach to measure ethnicity in Peru, showing our findings for the newly created data. Section 3 shows an analysis of the interplay among race and ethnic variables with a set of characteristics related to the performance of individuals in the labor market. Section 4 is devoted to the extension of the Blinder-Oaxaca decomposition to a setup in which there are not only two comparing groups but a continuum of them. Section 5 shows the main findings of the coefficient decompositions. Finally, conclusions are presented in Section 6.

2   Measuring Ethnicity

The anthropology literature, defines the concept of *ethnicity* in a general way as the community of individuals that share cultural elements and that organize their daily life around it. Generally, it is associated with the idea of native communities that are isolated or that have reduced contact with other communities. In urban settings, ethnic characteristics are associated, in a complex and passionately debated interaction,

---

[2] See Angel and Gronfein (1998) and Anderson, Silver and Abramson (1998) for previous use of this methodology.

with culture, religion, language, traditions and race, among other dimensions.

As mentioned above, we use information on mother tongue, religion and parental background to approximate ethnic differences. A more complex issue is the use of race indicators as an additional dimension of ethnicity. Several disciplines debate the complex interplay that exists between race and ethnicity. Here we just recognize that race, together with other ethnic characteristics, may generate differences among people that may have measurable consequences over economic opportunities. In order to approximate racial characteristics a score-based procedure is used, in which each individual received a score for each of the four categories: Asiatic, White, Indigenous, and Black; which are groups that are more easily recognized by people as distinct racial groups. Scores were given by each individual and independently by the pollster. This score ranges from 0 to 10 in each category; with zero meaning that the individual did not have physical characteristics that resembled a typical individual of the respective racial group and 10 that she had mostly characteristic features of that group. With this multi-dimensional racial intensity indicator, we are able to characterize a person as a Mestizo, but within those Mestizos, there is still racial variability.[3]

The self-report of race has been used in other countries with some success due to the fact that the classification of races in other places tends to be easier or more direct (see Hirshman and Alba, 1998 and Telles and Lim, 1998). However, in Peru, the majority of the population tends to define themselves as "Mestiza", category that includes people who actually have very different characteristics and are perceived by the others also as very different. The use of a second source of information, as it is the pollsters' perception is a technique supported by arguments as those exposed by Angel and Gronfein (1988) and Anderson, Silver and Abramson (1998). Even though this method might be also subjected to various criticisms, we tried to reduce the problems associated to the inter-observer variability with an intensive pre-fieldwork training, as suggested by Boergerhoff-Mulder and Caro (1985).[4]

There are three additional aspects through which we capture ethnic characteristics. First, we use language, a variable that typically has been used as the sole indicator of ethnicity in many labor studies in Latin America. Here we use language information for the individual and for his/ her parents. The second aspect is migration. In the data set, we consider both short-term migration as well as migration from place of origin, which is important due to the migratory process that has taken place in Peru in the last 50 years. Finally, since the religion of the individual might be relevant, we consider this variable as well.

---

[3] For certain econometric procedures, it was also useful to dichotomize our measure in order to identify three different groups: "Indigenous", "Whites" and "Mestizos" (Also Asian and Black were identified but the sample sizes were too small).

[4] In addition, following Allport and Kramer (1946), Scodel and Austrin, (1957), and Toch, Rabin and Wilkins (1962); we used pictures as instruments to standardize the pollster's reports. We carried out an intensive training in order to minimize inter-observer variability, homogenizing scoring criteria among all pollsters.

2.1 The Data

The data used comes from the urban households of the LSMS survey for 2000 and from an additional module carried out by GRADE in 2001. The latter was applied to a significant fraction of urban household members eighteen years or older and was designed to explore in depth racial and ethnic characteristics.[5] A novel and interesting feature of this additional module is the way the "race" variable was surveyed. As previously mentioned, the race of each individual had been considered as a four-dimensional vector (White, Indigenous, Black and Asian) with an ordinal measure of intensity, ranging from 0 (lowest) to 10 (highest), in each dimension independently.[6] For example, a predominantly white individual could be one with intensities 7, 1, 0 and 1 for the categories White, Indigenous, Black and Asian, respectively. An example of a predominantly indigenous individual could be one with intensities 2, 8, 0 and 1 in the same dimensions. With this feature, we are able to capture better the racial diversity and the different degrees of "mestizaje" present in the Peruvian society.

Figure 1 shows the distribution of the population according to the intensities in each race category, using the perceptions of the pollsters. As observed, the White intensity distribution is skewed towards the right, suggesting that the majority of individuals are characterized by pollsters as having some white characteristics, but are not predominantly white; on the other hand the indigenous intensity distribution is more skewed towards the left. Because populations with a strong Asian or Black ancestry are relatively small, the LSMS is not necessarily representative of these groups. Still, a small number of individuals have, according to pollsters' perceptions, racial features that resemble observable characteristics of these groups.

An immediate implication of the distribution of the population according to racial intensities will be our inability to establish statistical regularities regarding these latter groups. Therefore, we will concentrate the discussion on the consequences on earnings of racial differences among people with white and indigenous traits. In some statistical application where it is informative to divide the sample into groups, we use predominantly White, predominantly Indigenous and Mestizo as analytical categories.

As it is documented in the literature and shown in 2, there are significant differences between the race variable self-reported by the individuals surveyed and the same variable reported by the pollster (for example see Tellez (1998) for the Brazilian case).

---

[5] The module covered 70% of the original people surveyed (5,700 individuals). The lost by attrition of 30% of the cases did not represent significant differences in the main individual and household characteristics between this sub-sample and the total sampled population. In addition, the special module elaborated for this project included questions regarding physical characteristics, linguistic uses, geographic origin, religious habits, and information related to parents (mother tongue, geographic origin, religion and education). The survey also included a section dedicated to questions about credit, access to social capital of the people surveyed and cultural consumption. Finally, questions related to discrimination episodes were included.

[6] We did not impose ex-ante any condition on the values that the 4-dimensional vectors of racial characteristics may attain. That is, we left room for any of the 11x11x11x11=114 possible combinations of race intensities. Interestingly, the results obtained show low variability in terms of the sum of the four racial intensities for each individual.

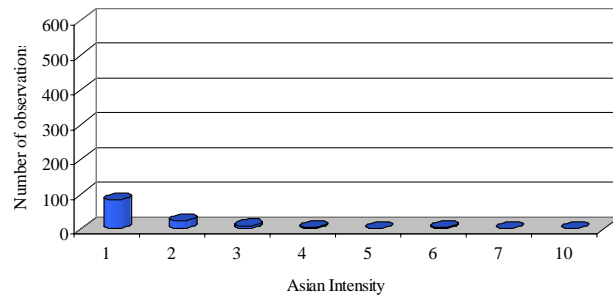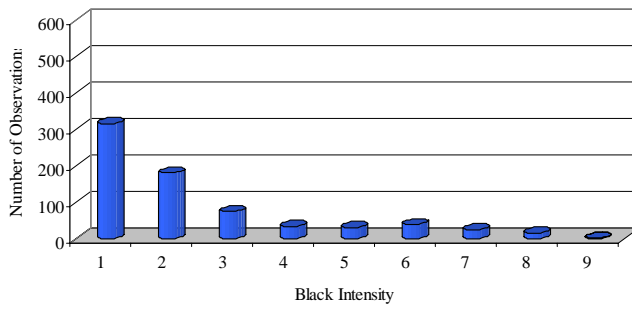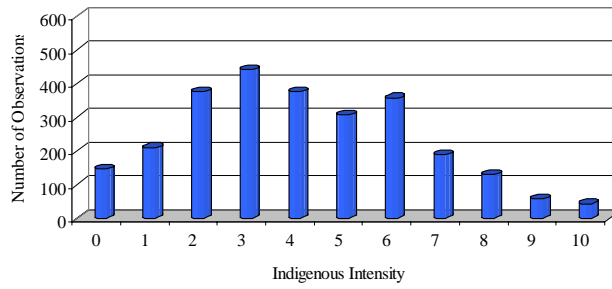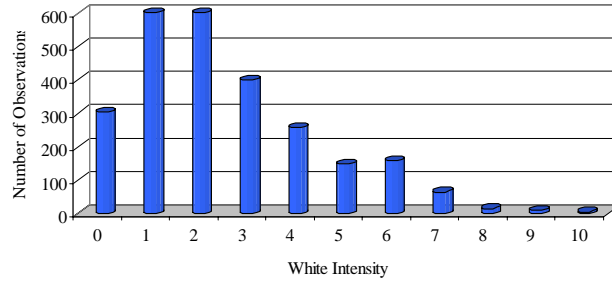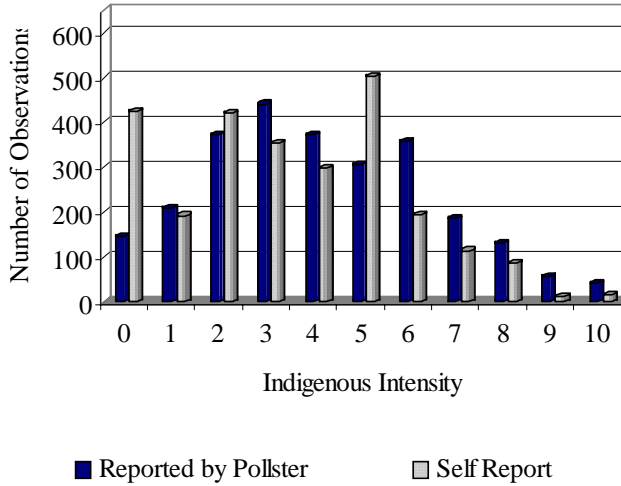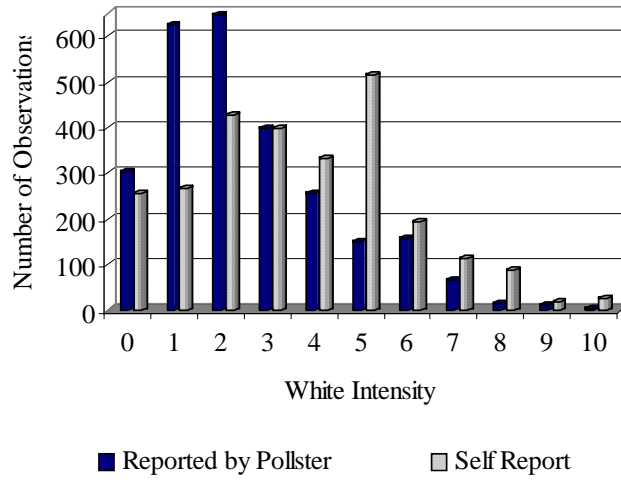Figure 1: Racial Intensity Distributions in Urban Peru - Pollster's Perception

Figure 2: Comparison between Race Intensity as Reported by Pollster and Self Reported

The self-reported White intensity distribution is skewed to the right of the same distribution reported by the trained pollster, while the self-reported indigenous intensity distribution is skewed to the left of the same distribution reported by the pollster. Overall, respondents tend to score themselves with higher values of white intensity and lower values of indigenous intensities than pollsters. That is, individuals consider themselves "less indigenous" than they are actually perceived by pollsters.

Given that the paper's main objective is to identify labor market exclusion given observable ethnic characteristics and not self-exclusion, we will concentrate on the pollster scores rather than the self reported scores. However, differences between both types of scores could have strong implications on the quantification of the racial earnings gaps.[7]

We are relying on the fact that people associate the words White, Indigenous, Black and Asian with different sets of phenotypic characteristics. Given other individual traits, these characteristics as perceived by a third person may or may not be associated with other socioeconomic variables or outcomes. Race, together with other ethnic-related characteristics, may have real effects in the labor markets, which we will try to approximate here, though we do not dwell into the specific sociological or economic mechanisms that may cause these effects.

3   Ethnicity and Individual Characteristics

Figure 3 present the relationship between racial intensities and some key variables.[8] It should be noted that the pollster records her own perception about racial intensities independently of the answers that the respondent gives about her characteristics.[9] Years of schooling are positively correlated with the White intensity indicator and negatively correlated with the Indigenous intensity. Similarly, the same pattern is observed regarding attending a private educational institution, access to phone lines and access to health insurance.

Individuals perceived as predominantly White report higher levels of education, a smaller family size and fewer children. Individuals that are predominantly Indigenous are less educated and have more children. Individuals perceived as having more Indigenous features report more frequently that their mother tongue is a native language, and are slightly more likely to be Christian non catholic, and are much more likely to be migrants. Regarding parental background, as the White intensity increases, mother's education is higher and the likelihood of the mother having a native language as a mother tongue is lower. Moreover,

---

[7]As it is shown by Telles and Lim (1998) for the Brazilian case, while the White-Brown gap is around 26% using the pollster perception, it is reduced to 17% if the self-report is used (both gaps are calculated controlling for human capital and labor market characteristics).

[8]We do not report White intensities of 9 and 10 as the number of observations for these cells is too small.

[9]Moreover, the main LSMS survey was applied a few months before people were re-interviewed for the additional ethnic module.

when we look at labor-related characteristics as the log of hourly income, the quintiles with higher White intensity had higher incomes than the quintiles with higher Indigenous intensities. Also, among the whiter quintiles there are more professionals and technicians, as well as executive staff.[10]

Figure 4 shows the relationship between racial intensities and poverty. The horizontal axis indicates the race intensities as perceived by the pollster and the vertical axis shows the proportion of poor individuals. It is clear that the higher the intensity of White the less poor are households and the higher the intensity of Indigenous the poorer they are.

Analyzing raw averages for the self-employed and private wage earners, the log hourly wage is positively correlated with the White intensity indicator and negatively correlated with the Indigenous intensity. In both cases and at the same time the average levels of earnings are lower for the self-employed than for the private wage earners (see Figure 5 and Figure 6).

Mother tongue seems to have a strong correlation with raw earnings but only among the self-employed. Also, among the self-employed there is a small difference in earnings depending on migrant status. There are differences in log hourly wages between workers of different religions. Finally, being born in a rural area is correlated with lower earnings, independent of the type of job, as shown in Table 2. All these variables, however, may be correlated with human capital and demographic variables, so further analysis is required, as presented in the next two sections.

4   An Extension of the Blinder-Oaxaca Decomposition to a Continuum of Comparisons.

The Blinder-Oaxaca approach to decompose the wage gap between two comparing groups can be naturally extended into a setup in which there are not only two groups but a continuum of them (as it is the case with the racial characteristics of the Peruvian population). This section is devoted to introduce such extension. With the purpose of simplifying the presentation, we will change slightly the traditional notation involved in the Blinder-Oaxaca decomposition. After this, the extension to a continuous setup will seem natural.

In a setup in which there are two comparing groups (males and females, blacks and whites, formal and informal workers, etc.), let the dummy variable t_i indicate the group at which individual i belongs ($t_i = 0$ for those individuals i that belong to the base group and $t_i = 1$ for those individuals i that belong to the comparing group). Denoting by y_i the earnings of individual i, the wage gap can be expressed as

$$\alpha_1 = E[y|t=1] - E[y|t=0] \tag{1}$$

---

[10]Detailed tabulations of these demographic characteristics are available from the authors upon request.

Figure 3: Selected Individual and Family Characteristics by Racial Intensity

Years of Schooling by Racial Intensity

Studies in a Private Institution by Racial Intensity

Access to Health Insurance by Racial Intensity
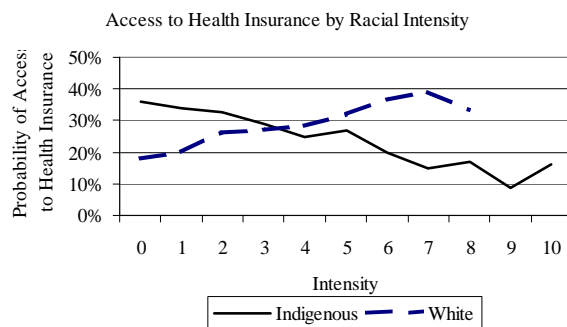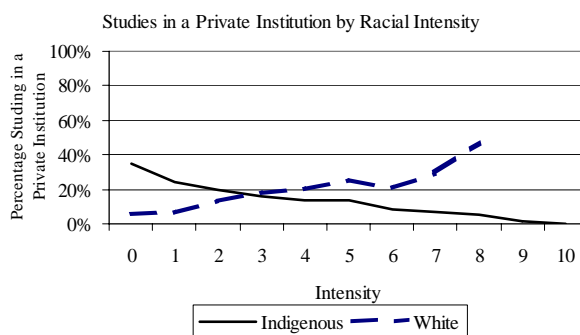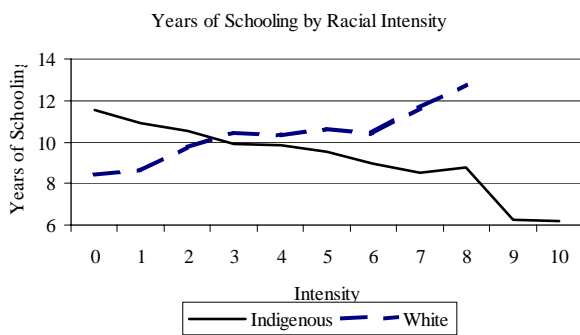
Figure 4: Poverty Index by Racial Intensity
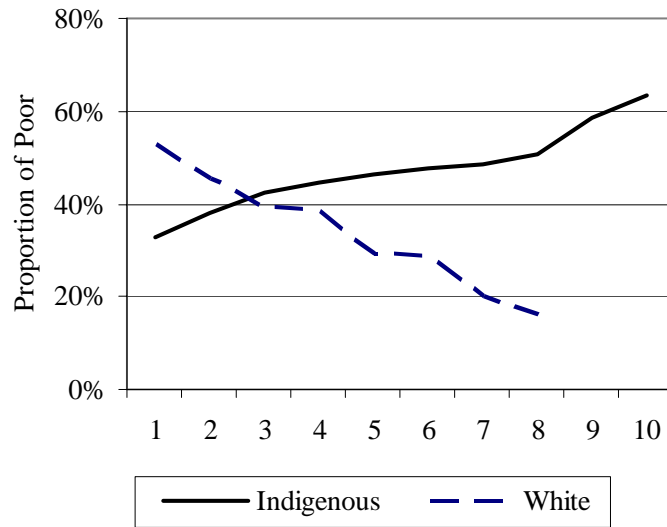
Figure 5: Hourly Earnings by Racial Intensity and Type of Employment
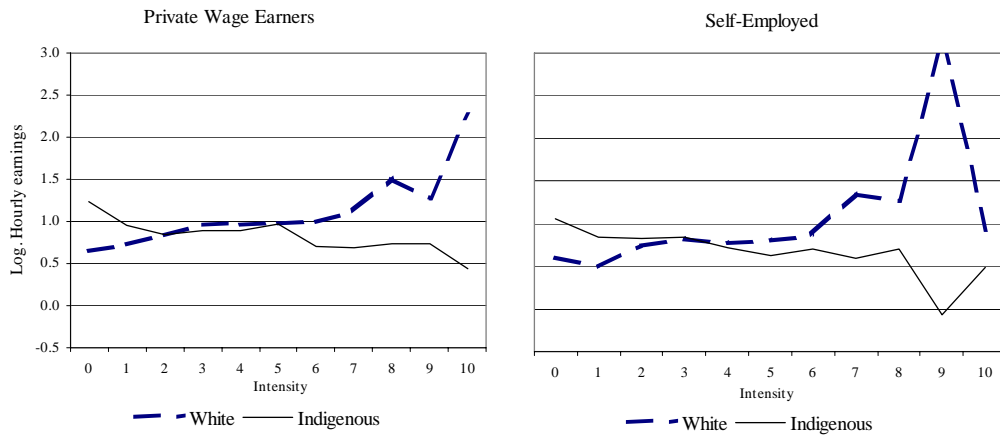
Figure 6: Log hourly earnings by other ethnic characteristics and type of employment

		Private Wage Earners	Self Employed
Average of Log hourly earnings		0.89	0.7
Mother Tongue			
	Spanish	0.87	0.76
	Native language	0.82	0.46
Religion			
	Catholic	0.88	0.74
	Christian non catholic	0.8	0.53
	Other religions	1.02	0.45
	No religion	0.85	1.08
Migrant Status			
	Migrant	0.88	0.76
	Non migrant	0.87	0.69
Birthplace			
	Born in rural or semirural area	0.67	0.33
	Born in an urban area	0.89	0.79

Which in turn can be computed from the estimated slope of the equation

$$y_i = \alpha_0 + \alpha_1 t_i + \varepsilon_i \tag{2}$$

In this setup, the Blinder-Oaxaca approach decomposes the α_1 coefficient, the wage gap, on the basis of observable characteristics. For that purpose, it is necessary to estimate the (Mincerian) earnings equation

$$y_i = \beta_0 + \beta_1 x_i + \beta_2 t_i + \beta_3 x_i t_i + \nu_i \tag{3}$$

where t is the dummy variable introduced above, x is a n-dimensional vector of observable characteristics (age, education and occupational experience, among others), β_1 and β_3 are the corresponding n-dimensional vectors of "rewards" for those characteristics (β_1 for individuals of group 0 and $\beta_1 + \beta_3$ for individuals of group 1) and β_0 and β_2 are one-dimensional coefficients. Then, 1 can be expressed as

$$\alpha_1 = E\left[(\beta_0 + \beta_2) + (\beta_1 + \beta_3)\,x|t = 1\right] - E\left[\beta_0 + \beta_1 x|t = 0\right]$$

which after some re-arrangements becomes

$$\alpha_1 = (\beta_0 + \beta_2) + (\beta_1 + \beta_3)\,E\left[x|t = 1\right] - \beta_0 - \beta_1 E\left[x|t = 0\right]$$

$$\alpha_1 = \beta_1\left(E\left[x|t = 1\right] - E\left[x|t = 0\right]\right) + \beta_2 + \beta_3 E\left[x|t = 1\right]$$

$$\alpha_1 = \Delta_x + \Delta_0. \tag{4}$$

Now, $\Delta_x \equiv \beta_1\left(E\left[x|t = 1\right] - E\left[x|t = 0\right]\right)$ and $\Delta_0 \equiv \beta_2 + \beta_3 E\left[x|t = 1\right]$ can be interpreted in the traditional way. Δ_x is the component of the wage gap that is explained by the difference in average characteristics of the individuals, while Δ_0 is the component that remains unexplained and can be attributed to the existence of a combination of discrimination and unobservable characteristics.

This setup in which t is a binary variable can be extended into another in which t is continuous. Now, 1 from equation 2 has to be rewritten as

$$\alpha_1 = \frac{cov\left(y, t\right)}{var\left(t\right)}.$$

Just for presentation purposes, without loss of generality, let us assume $E\left(t\right) = 0$ and $E\left(t^2\right) = 1$,[11] then

---

[11] As I will discuss later in this section, the two moment assumptions made here have the only purpose of algebra simplification. The results are still valid without any of them.

$\alpha_1 = E[yt]$ or

$$\alpha_1 = \int_D tE[y|t]\, dF(t),$$

where D represents the domain and F represents the cummulative distribution for t. Next, using the earnings equation 3 in the new expression for α_1 will lead to

$$\alpha_1 = \int_D (\beta_0 + \beta_1 E[x|t] + \beta_2 t + \beta_3 tE[x|t])\, t\, dF(t).$$

Using the properties assumed on the first two moments of t we have

$$\alpha_1 = \int_D (\beta_1 + \beta_3 t)\, E[x|t]\, t\, dF(t) + \beta_2. \tag{5}$$

Now, noting that $\beta_3 E[x|t = t_0]$ is equivalent to $\int_D (\beta_1 + \beta_3 t)\, E[x|t = t_0]\, t\, dF(t),$ we can add the former and subtract the latter from the right-hand side of 5 to get, after some re-arrangements

$$\alpha_1 = \left[\int_D (\beta_1 + \beta_3 t)\, (E[x|t] - E[x|t = t_0])\, t\, dF(t) \right] + [\beta_2 + \beta_3 E[x|t = t_0]]$$

$$\alpha_1 = \widetilde{\Delta_x} + \widetilde{\Delta_0}. \tag{6}$$

Which now has an analogous interpretation to the Blinder-Oaxaca decomposition.

$$\widetilde{\Delta_x} = \int_D (\beta_1 + \beta_3 t)\, (E[x|t] - E[x|t = t_0])\, t\, dF(t)$$

is the component of α_1 that can be explained by the aggregated differences in average characteristics between individuals of type t_0 and the rest of the population.

$$\widetilde{\Delta_0} = \beta_2 + \beta_3 E[x|t = t_0]$$

is the component of α_1 that remains unexplained.

The two moment conditions about t were assumed only for expositional purposes. Without imposing any of those assumptions on t, the decomposition changes only slightly, it becomes

$$\widetilde{\Delta_x} = \frac{\int_D (\beta_1 + \beta_3 t)\, (E[x|t] - E[x|t = t_0])\, (t - E[t])\, dF(t)}{var(t)}$$

and

$$\widetilde{\Delta_0} = \beta_2 + \beta_3 E\left[x|t = t_0\right].$$

Note that the expression for the unexplained component, $\widetilde{\Delta_0}$, as a function of β_2, β_3 and $E\left[x|t = t_0\right]$, does not change.

Also, instead of adding $\beta_3 E\left[x|t = t_0\right]$ and substracting its equivalent form, it is possible to simply add the unconditional version $\beta_3 E\left[x\right]$ and substract its equivalent form $\int_D \left(\beta_1 + \beta_3 t\right) E\left[x\right] t dF\left(t\right)$. In such a way the decomposition becomes:

$$\widetilde{\Delta_x} = \frac{\int_D \left(\beta_1 + \beta_3 t\right) \left(E\left[x|t\right] - E\left[x\right]\right) \left(t - E\left[t\right]\right) dF\left(t\right)}{var\left(t\right)}.$$

and

$$\widetilde{\Delta_0} = \beta_2 + \beta_3 E\left[x\right].$$

Which can be interpreted in the light of a slightly different counterfactual situation, analogous to the one in [?].

Finally, it is straightforward to verify that for the particular case in which t is binary and $t_0 = 1$ (the traditional Blinder-Oaxaca setup), $\widetilde{\Delta_0}$ in (4) becomes equal to Δ_0 in (6). The decomposition I introduced in this section is a generalization of the traditional Blinder-Oaxaca decomposition for a situation in which, instead of two comparing groups, there are infinitely many of them.

5 Racial Wage Gaps in Urban Peru

For the setup of this paper, the continuous variable t will be approximated by a variable p such that p_i denotes the percentile of individual i in a ordering of race intensities where the primary ordering key is the white intensity (treated in an ascending order) and the secondary ordering key is the indigenous intensity (treated in a descending order). Then, we can consider p as a continuous variable in the interval $[0,1]$. The base-group in this application is conformed by those individuals in the lowest percentiles of the white distribution, that is $t_0 = 0$.

We estimated earnigs equations, controlling for sex, age (and its square), years of schooling, marital status, years of occupational experience (and its square), occupation (eight dummies for nine occupational groups), firm size (four dummies), city (a dummy that distinguishes Lima from the rest of the nation), mother's educational level (two dummies), number of sick days during the last year, migratory condition (a dummy) and hours worked per week. The results obtained for an estimation with the whole national sample are $\alpha_1 = 0.4947$ and $\widetilde{\Delta_0} = 0.1195$. According to these figures, the average individual with the

Figure 7: Racial Coefficient Decomposition by Different Partition Criteria

|	Coefficient Alpha-1	Unexplained Component
All	0.4947	0.1195
Geographic		
Lima	0.5658	0.2448
Rest of the Nation	0.3427	0.1039
Gender		
Females	0.6466	0.2192
Males	0.4105	0.0751
Type of Employment		
Private Wage Earners	0.4989	0.1986
Public Wage Earners	0.2815	0.1794
Self-Employed	0.522	0.1086

highest white intensity (percentile 100) earns approximately 49.47% more than the average individual with the lowest white intensity (percentile 1) in per hour terms. This difference is partially explained by the fact that the average individual with the highest white intensity has individual characteristics that exceeds those of the average individual with the lowest white intensity. After accounting for those differences in observable characteristics, there is still a difference of 11.95% in favor of the white individulas. A confidence interval for $\widetilde{\Delta_0}$, the measure of unexplained differences in wages, ranges from 0.5% to 23.4%.[12]

Partitioning the sample by different criteria will shed some lights. Next, in Table 7 we present the coefficients for averages differences in wages (α_1) and their corresponding unexplained components $\left(\widetilde{\Delta_0}\right)$ for partitions that consider geographic, gender and type of employment criteria. The results suggests that there are more evidences of differences in pay that cannot be explained by differences in observable characteristics among those who live in Lima, those who are females and those who work as dependants (either in the private sector or in the public sector).

A simultaneous look at gender and racial differences in earnings reveals that the gender wage gap varies with racial characteristics, being the case that the gap is smaller among those with white characteristics. If the gap among males and females at the lowest percentile of white intensity is around 30.71%, it attains 16.30% between those males and females at the highest percentile (after accounting for differences in observable characteristics), as we can see in Figure 8.

---

[12]This result has been obtained from 5000 bootstrap iterations.

Figure 8: Average Wages By Gender and Race

Average Wages By Percentiles of White Intensity
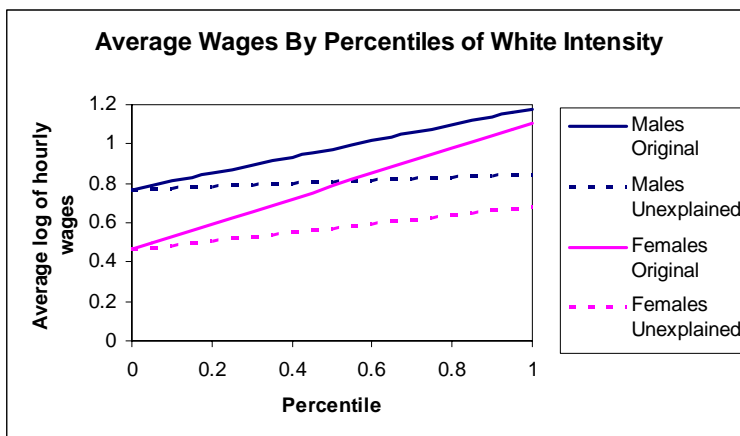
6   Conclusions

Given that Peru is an extremely diverse country where ethnic groups cannot be easily identified, we approximate the ethnic diversity of the country using a large set of variables like language, religion, origin and race. The use of race indicators in a country like Peru is complex. Here we recognize that race, together with other ethnic characteristics, may generate differences among people that may have measurable consequences with regard to economic opportunities. In order to approximate racial characteristics we use a score-based procedure in which each individual received a score for each of the four categories Asiatic, White, Indigenous, and Black, which are groups that are more easily recognized by people as distinct racial groups. Scores were self-reported by each individual and were also assigned, independently, by the pollster. With this multi-dimensional racial intensity indicator, we are able to characterize a person as a Mestizo, but within those Mestizos, there is still racial variability that we can capture and explore.

The self-reported White intensity distribution is skewed to the right of the same distribution reported by the trained pollster, while the self-reported Indigenous intensity distribution is skewed to the left of the same distribution reported by the pollster. Overall, respondents tend to score themselves with higher values of White intensity and lower values of Indigenous intensities than pollsters score them. The empirical analysis shows that our race indicators are clearly related to poverty variables and specific assets. For instance, individuals who have higher intensities in the White scale have a lower poverty index, more years of schooling, greater access to phone lines, and more access to health insurance and to private education.

Extending the traditional decomposition of wage gaps (Blinder-Oaxaca) to a setup in which, instead of comparing two groups, there is a continuous of them, we found interesting results. In Urban Peru, the

17

extremely white individuals earn approxiamtely 50% more than the extremely indigneous. After accounting for the differences in other observable characteristics related to earnings, there is still a premium of 12%.

The results suggest that there is more evidence of racial discrimination in Lima than in the rest of the nation, also more discrimination among females than among males, and finally, more evidence of discrimination among the wage earners (public or private) than among the self-employed.

There is an interesting link between racial and gender differences in earnings. The gender wage gap is bigger among those individuals with the fewest white observable characteristics.

References

[1] Allport, Gordon, and Bernard Kramer; (1946). "Some roots of prejudice." The Journal of Psychology, 22 (1946), pp.9-39.

[2] Anderson Barbara A., Brian D. Silver and Paul R. Abramson; (1988). "The effects of the Race of the Interviewer on Race-Related Attitudes of Black Respondents in SRC/CS National Election Studies." Public Opinion Quarterly, Vol 52, No. 3, pp. 289-324.

[3] Angel, Ronald and William Gronfein; (1988). "The Use of Subjective Information in Statistical Models." American Sociological Review, Vol. 53, pp. 464-473.

[4] Arce, C.H., E. Murguia and W.P. Frisbie; (1987). "Phenotype and Life Chances Among Chicanos." Hispanic Journal of Behavioral Sciences, 9:19-32.

[5] Barsky Robert, John Bound, Kerwin Charles and Joseph Lupton (2001). "Accounting for the Black-White Wealth Gap: A Nonparametric Approach." *NBER Working Paper* No. 8466, National Bureau of Economic Research, Cambridge, MA.

[6] Becker, Gary; (1971). "The Economics of Discrimination." Second Edition. Chicago: The University of Chicago (1971).

[7] Blinder, A.; (1973). "Wage Discrimination: Reduced Form and Structural Estimates." The Journal of Human Resources, VII, 4, pp. 436-455.

[8] Boergerhoff-Mulder and T. M. Caro; (1985). "The Use of Quantitative Observational Techniques in Anthropology." Current Anthropology, Vol. 26, pp. 323-335.

[9] Callirgos, Juan Carlos; (1993). "El racismo: la cuestión del otro (y de uno)." – Lima : DESCO, 235 p.

[10] Chiswick, Barry; (1988). "Differences in Education and Earnings across Racial and Ethnic Groups: Tasetes, Discrimination, and Investment in Child Quality." The Quarterly Journal of Economics, 103(3): 571-597, 1988.

[11] Darity, William and Patrick Mason; (1998). "Evidence on Discrimination in Employment: Codes of Color, Codes of Gender." The Journal of Economics Perspectives, 12:63-90.

[12] Deutsch, Ruthanne, Andrew Morrison, Claudia Piras and Hugo Ñopo (2002). "Working Within Confines: Occupational Segregation by Gender in Three Latin American Countries." *Inter-American Development Bank. Sustainable Development Department. Technical Papers Series No. 126.* Washington D.C.

[13] DiNardo, John; Nicole Fortin and Thomas Lemieux (1996). "Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach." *Econometrica*, Volume 64, No. 5, 1001-1044.

[14] Efron, B. and Tibshirani, R.; (1991). "Statistical Data Analysis in the Computer Age" In Science, 1991, Vol. 253, pp. 390-395.

[15] Efron, B. and Tibshirani, R.; (1993). "An Introduction to the Bootstrap" (New York: Chapman & Hall).

[16] Hirshmann, Charles and Richard Alba; (1998). "The Meaning and Measurement of Race in the U.S. Census: Glimpses into the Future." Paper presented at The Annual Meeting of the American Sociological Association. 1998.

[17] Horowitz, Joel; (2001). "The Bootstrap." In Handbook of Econometrics, Vol. 5, J.J. Heckman and E. Leamer, eds., Elsevier Science B.V., 2001, pp. 3159-3228.

[18] Johnson, J.H. and W.C. Farrell; (1995). "Race Still Matters." In Chronicle of Higher Education, Vol XLI, No. 43 (July 7),: A48.

[19] Longhofer, Stanley D. and Stephen Peters (1998). "Self-Selection and Discrimination in Credit Markets." Federal Reserve Bank of Cleveland. Originally published July 1998. Revised June 2000.

[20] MacIsaac, Donna; (1993). "Peru." In Indigenous People and Poverty in Latin America: An Empirical Analysis. George Psacharopoulos and Harry Patrinos. (Ed) The World Bank. Latin American and the Caribbean technical Departments. Regional Studies Program. Report No. 30. pages 179 - 224.

[21] Mendoza Arroyo, Amalia, ed.; (1993). "Cuestión de piel: testimonios de racismo en el Peru." Lima: ADEC-ATC, 1993. 132 p.

[22] Ñopo, Hugo (2002). "Matching as a Tool to Decompose Wage Gaps". Unpublished. Northwestern University.

[23] Oliart, Patricia; (1989). "El cristal con que se mira. Algo sobre género, raza y clase en el Peru." Lima : PUCP, 1989. Mimeo. 20 p.

[24] Oaxaca, Ronald; (1973). "Male-Female Wage Differential in Urban Labor Market." International Economic Review, Vol. 14, No. 3, (Oct. 1973), pp. 693-709.

[25] Oaxaca, R. & M. Ransom (1994). "On Discrimination and the Decomposition of Wage Differentials." *Journal of Econometrics,* 61 (1), 5-21.

[26] Pozzi-Escott, Ines. "La Discriminación Etnico Cultural en la Escuela Peruana". Lima: P.U.C.P. 1989.

[27] Psacharopoulos, George and Harry Patrinos, eds.; (1994). Indigenous Peoples and Poverty in Latin America: An Empirical Analysis. Washington, DC: The World Bank.

[28] Robinson, P.M.; (1988). "Root-n-consistent Semiparametric Regression." Econometrica, 56, 931-954.

[29] Saavedra, Jaime; (1997). "¿ Quienes Ganan y Quienes Pierden con una reforma estructural?: cambios en la distribución en la dispersión de ingresos según educación, experiencia y genero en el Peru urbano." Notas para el debate No. 14. GRADE. Lima., August 1997, pages 9-77.

[30] Scalera, Domenico and Alberto Zazzaro; (2001). "Group reputation and persistent (or permanent) discrimination in credit markets." Journal of Multinational Financial Management, Vol. 11 (2001) pp 483-496.

[31] Scodel, Alvin and Harvey Austrin; (1957). "The perception of Jewish photographs by Non-Jews and Jews." Journal of Abnormal and Social Psychology, 54, March, pp. 278-280.

[32] Silva, N.V.; (1985). "Updating the Cost of Not Being White in Brazil." In P.-M. Fontaine, ed., Race, Class and Power in Brazil. Los Angeles: University of California.

[33] Silva, N.V.; (1992). "Racial Discrimination in Income." Paper presented at the International Seminar, Labor Market Roots of Poverty and Inequality in Brazil. August 12-14, Rio de Janeiro.

[34] Sulmont Samain, Denis; (1995). "Exclusión social y empleo: notas para un debate"/ Denis Sulmont Samain.

[35] Telles, Edward and Nelson Lim; (1998). "Does it matter who answers the race question? Racial classification and income inequality in Brazil?." Demography, 35 (4). November

[36] Toch, Hans, Albert Rabin, and Donald Wilkins; (1962). "Factors Entering into Ethnic Identifications: an Experimental Study." Sociometry, 25, September 1962, pp.297-312.

[37] World Bank; (1999). "Peru, Education at a Crossroads: Challenges and Opportunities for the 21st Century". Human Development Department. LACR. Report No. 19066.

IZA Discussion Papers

No.	Author(s)	Title	Area	Date
965	M. Merz E. Yashiv	Labor and the Market Value of the Firm	1	12/03
966	T. Palokangas	Optimal Taxation with Capital Accumulation and Wage Bargaining	3	12/03
967	M. Lechner R. Vazquez-Alvarez	The Effect of Disability on Labour Market Outcomes in Germany: Evidence from Matching	6	12/03
968	M. Blázquez M. Jansen	Efficiency in a Matching Model with Heterogeneous Agents: Too Many Good or Bad Jobs?	1	12/03
969	J.-P. Schraepler G. G. Wagner	Identification, Characteristics and Impact of Faked Interviews in Surveys	7	12/03
970	G. Kertesi J. Köllő	Fighting "Low Equilibria" by Doubling the Minimum Wage? Hungary's Experiment	4	12/03
971	J. De Loecker J. Konings	Creative Destruction and Productivity Growth in an Emerging Economy: Evidence from Slovenian Manufacturing	4	12/03
972	J. Köllő	Transition on the Shop Floor - The Restructuring of a Weaving Mill, Hungary 1988-97	4	12/03
973	C. Belzil J. Hansen	Structural Estimates of the Intergenerational Education Correlation	1	12/03
974	J. Schwarze M. Härpfer	Are People Inequality Averse, and Do They Prefer Redistribution by the State? A Revised Version	3	12/03
975	A. Constant K. F. Zimmermann	Occupational Choice across Generations	1	12/03
976	J. D. Angrist K. Lang	Does School Integration Generate Peer Effects? Evidence from Boston's Metco Program	6	01/04
977	M. Corak G. Lipps J. Zhao	Family Income and Participation in Post-Secondary Education	5	01/04
978	J. T. Addison P. Portugal	How Does the Unemployment Insurance System Shape the Time Profile of Jobless Duration?	3	01/04
979	M. Moreno H. Ñopo J. Saavedra M. Torero	Gender and Racial Discrimination in Hiring: A Pseudo Audit Study for Three Selected Occupations in Metropolitan Lima	1	01/04
980	H. Ñopo J. Saavedra M. Torero	Ethnicity and Earnings in Urban Peru	1	01/04

An updated list of IZA Discussion Papers is available on the center's homepage www.iza.org.