



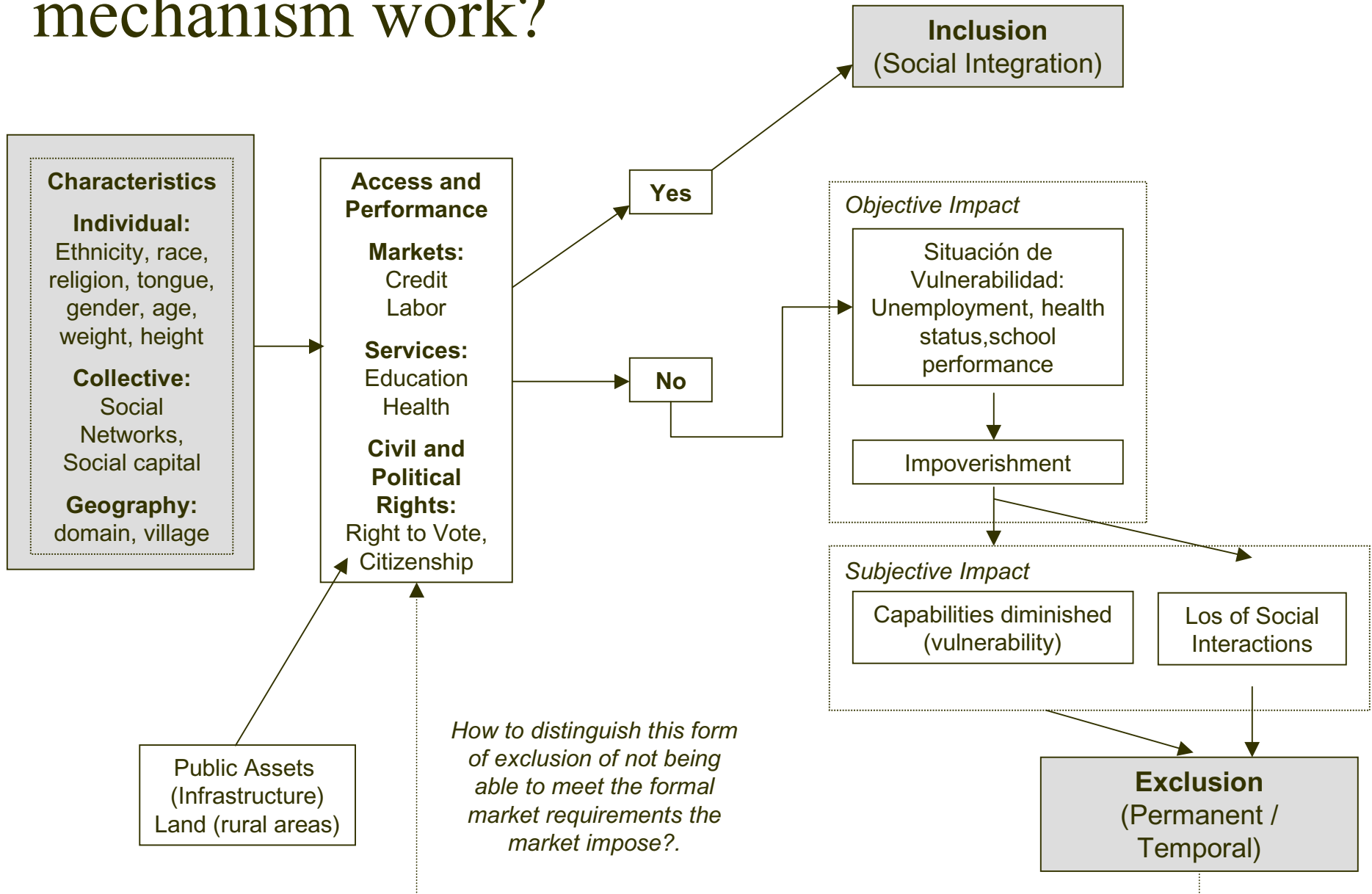
# Ethnicity and Earnings in Urban Peru

Hugo Ñopo, Jaime Saavedra and Máximo Torero  
Grupo de Análisis para el Desarrollo (GRADE)

October 2001

This paper is part of the project on Social Exclusion  
sponsored by the SDS/IDB

# How do the social exclusion mechanism work?



# About the concept

- Origin: France (Lenoir, 1974).
- Main approaches: french (solidarity), british (market failures) and nordic (power monopoly).
- Different emphasis on groups of analysis between Europe and Underdeveloped Countries: migrants, unemployed versus ethnics minorities/majorities o racial groups.
- A multidimensional phenomena that covers: economics, politics and culture. Acts as a complementary view of poverty and inequality.
- Although, it can not be directly “measured”, it might be proxied by some specific forms it may asume.

# About Labor Income (biblio)

**Tomes (85).** Dummies for religions (Jews, Protestants and Roman Catholics) and testing for differences in different Canadian regions.

**Borjas (87).** The endogeneity of the migration decision in analyzing the performance of migrants in the labor market.

**Chiswick (88).** Again, the racial and ethnic groups considered are based on nationalities. He looks at White (native born parents, foreign born parents, Jewish, Non-Jewish, Mexican origin), Black (Native born parents, Foreign born parents), Japanese, Chinese, Filipino, and American-Indian.

**Borjas (89).** With a theoretical model they find an equilibrium in which able Blacks are less likely to self-select into the self-employment sector rather than able Whites. Also, with 1980 Census data they confirm the model.

# About Labor Income (biblio)

**Farley (90).** Identifies ethnic groups with nationalities in USA. Based on the 1980 census he found differences between groups. His classification is like this: Hispanic groups (Cuban/Mexican/Puerto Rican/Other Spanish), Racial Minorities (Amer. Natives/Asian Indians/Black/Chinese/Filipino/Japanese/Korean/Vietnamese/Other races), White Ethnic Groups (American/American Indian /Armenian /Austrian /Belgian/ Canadian/ Croatian/ Czech/Danish /Ukranian /Yugoslavian/NEC).

**Borjas (92).** Ethnic skill differentials transmitted across generations. Using the self-reported ethnic classification of 1970 Census. And again, the ethnic groups are almost country groups.

**Craig (93).** Found occupational crowding for some groups (Blacks and Portuguese people) but little evidence of wage discrimination in the nineteenth century American whaling industry. For this purpose they used a set of race dummies in the earning equations

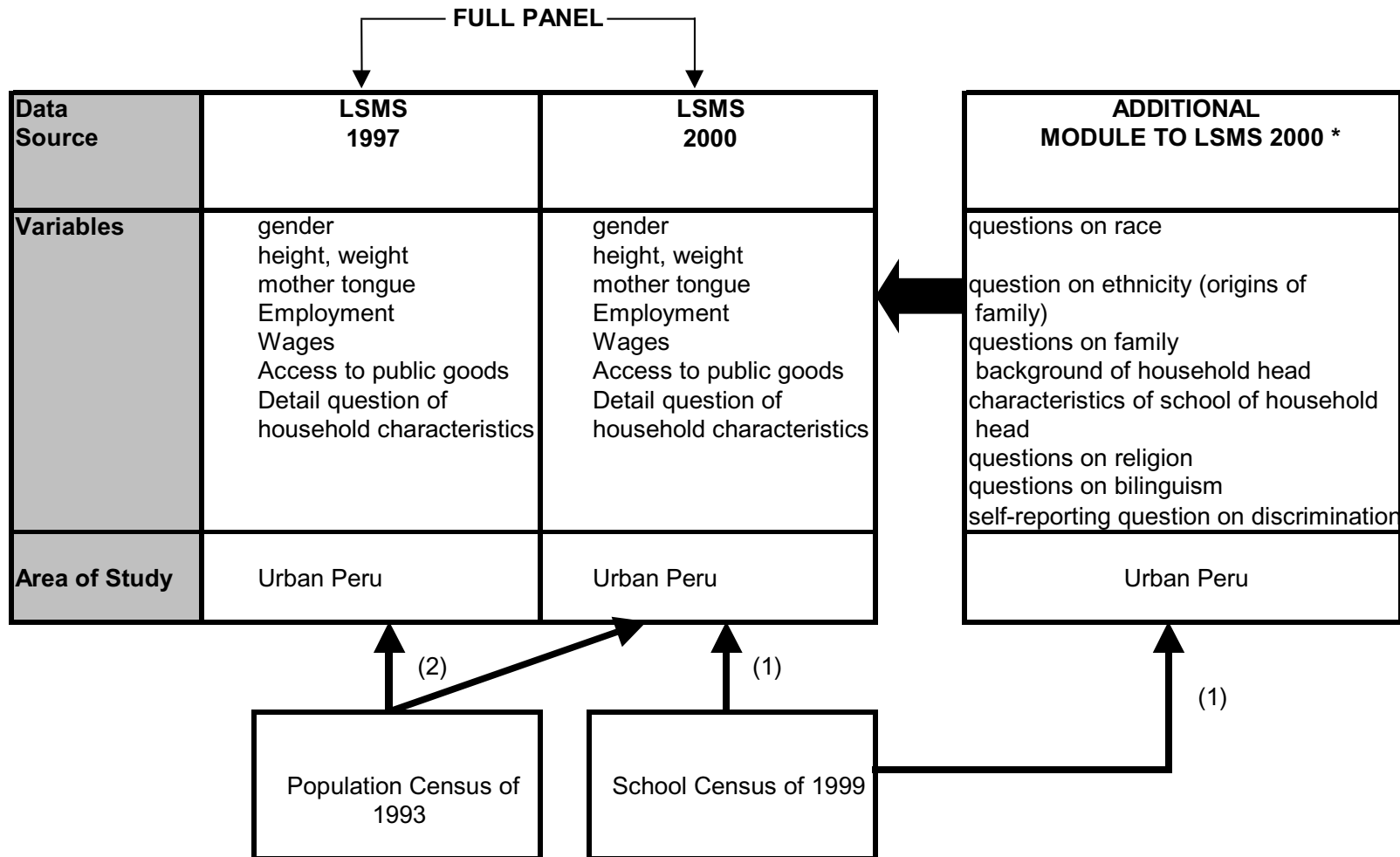
# About Labor Income (biblio)

**Sundstrom (94).** Occupational segregation, 18 cities in the U.S. 1910-1950. He computed Duncan indices and after computing it for many different cities he run a regression trying to explain the indices in terms of residential segregation and region.

**Bayard (99).** Greater segregation between Hispanic men and White men than between Hispanic women and White women accounts for almost all the higher Hispanic-White wage gap for men. For this purpose they use variables of the type: percentage of Hispanics in the occupation where the individual is. These types of variables typically have problems of correlation with the error (Usui(2001)).

**Gaviria (2001).** Cross countries comparison. Latin-America late nineties. Dummies for race (one for Blacks and one for Indigenous), very simple. Interesting thing: the survey (Latinobarometro) have questions about self-perception of discrimination in an ordinal scale (0 to 10) in two “interaction spaces” (jobs and schools) and three “organizations” (political parties, police, judiciary system).

# The Data



(1) Merged by School Name

(2) Merged by district name

# Information Collected in Additional Module to LSMS 2000

- Individual characteristics: ethnic group (proxied by place of birth and language), racial group (self reported and reported by pollster).
- Self reported physical characteristics: weight, height
- Linguistic history (mother tongue, language used at school, bilingualism)
- Information on Parents (place of birth, language, education, employment, racial group)
- Religion

# Social Exclusion.

## Ethnic Characteristic

- Racial Characteristics
  - Four dimensional vector with a ordinal measure of racial intensity in each one.
- Religion
- Mother's tongue at birth

# Race indicator construction

Race of each individual is characterized by a 4 dimensional vector, with an ordinal measure of race intensity from 0 to 10 in each dimension.

For example:

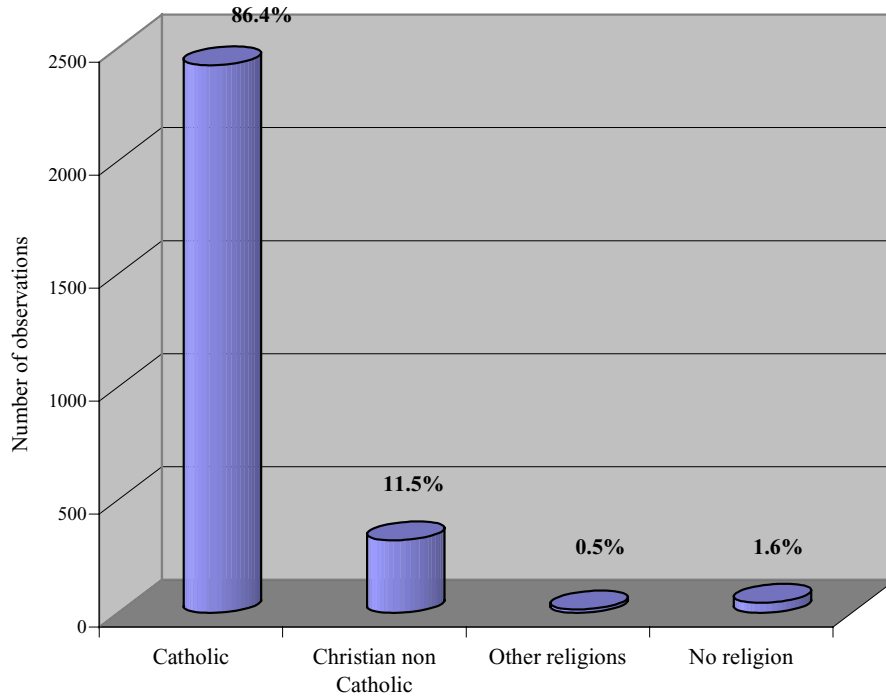
	Indigenous	White	Black	Asian
Predominantly White Individual	1	8	0	0
Predominantly Indigenous Individual	8	2	0	1
Typical "Mestizo"	3	4		0

# Caveats and other issues

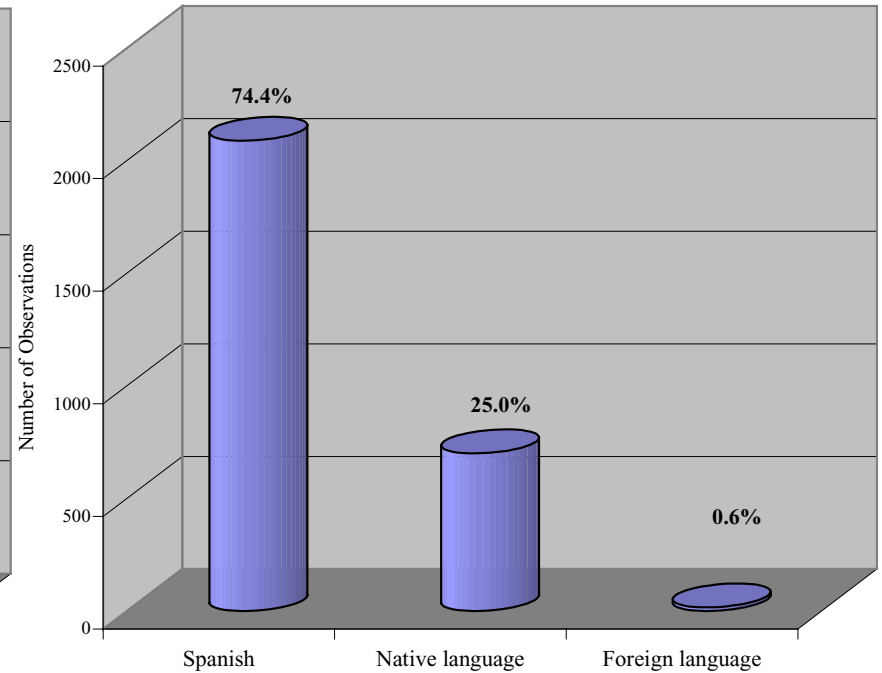
- Report is done by pollster and individual independently
- Pollster where trained in order to reduce sistematic differences in perceptions
- No restrictions were given to a maximum in the 0-10 ordinal scale for the 4 groups as a whole. However, “adding” the figures for each of the 4 race intensities never exceed 18
- Normalization may be required.

# Distribution of population by Religion and Mother Tongue

## Religion Distribution

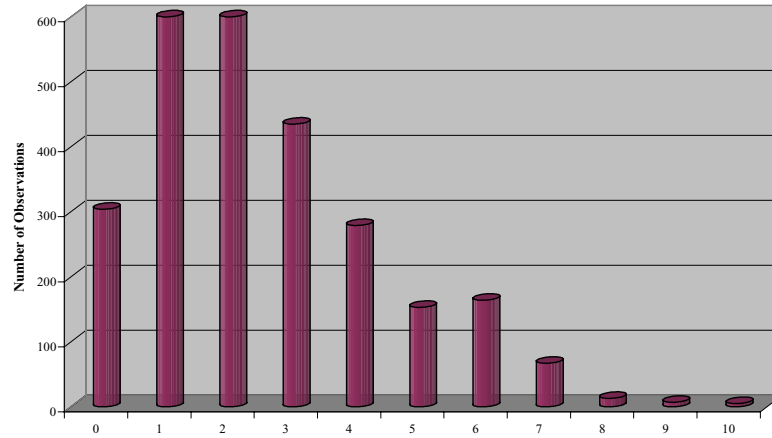


## Mother's Tongue Distribution

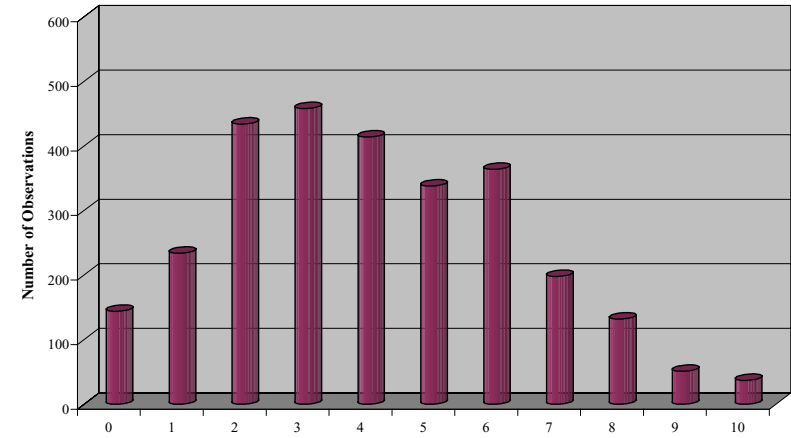


# Racial Characteristics

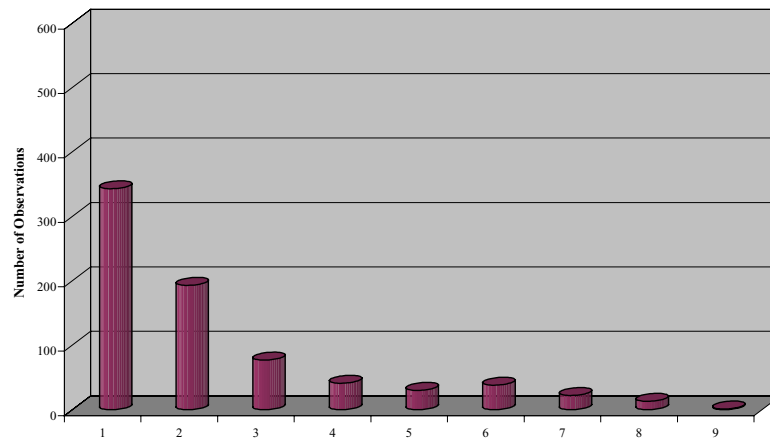
## White Intensity Distribution



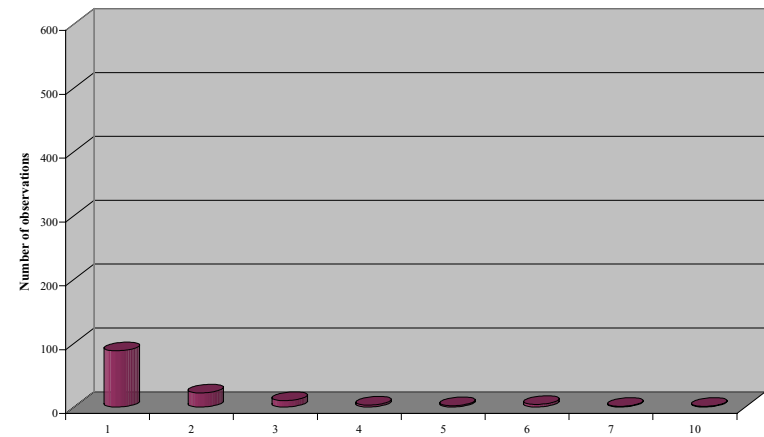
## Indigenous Intensity Distribution



## Black Intensity Distribution

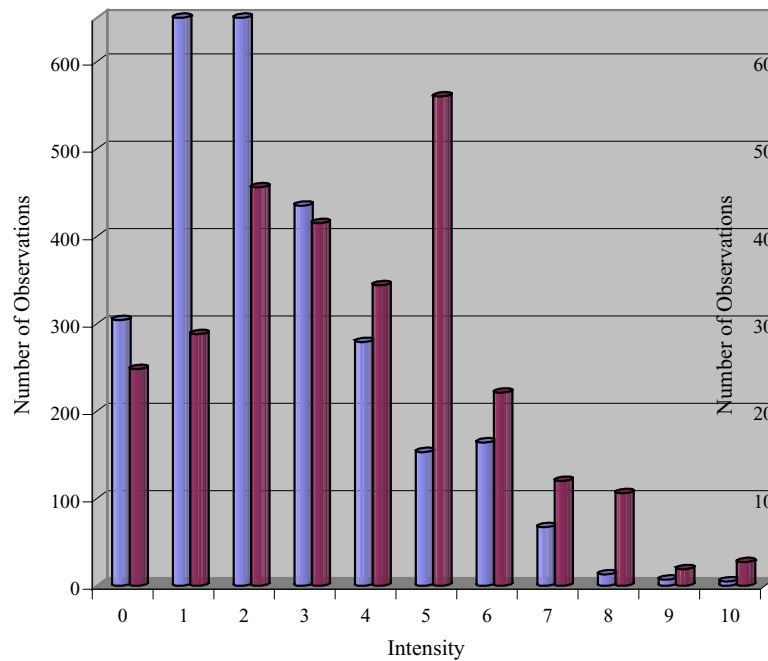


## Asian Intensity Distribution

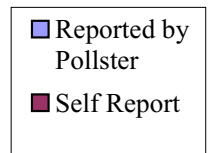
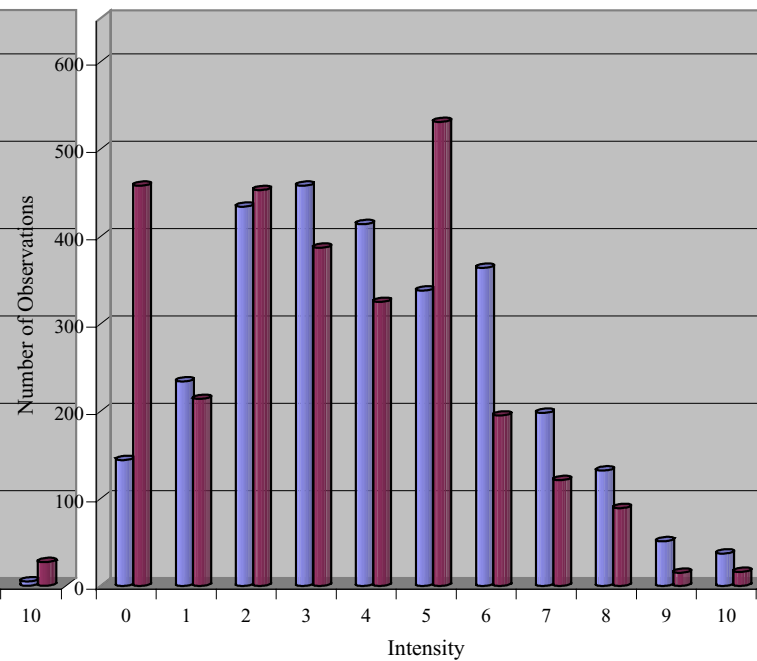


# Racial Characteristics: Self Report vs. Pollster Report

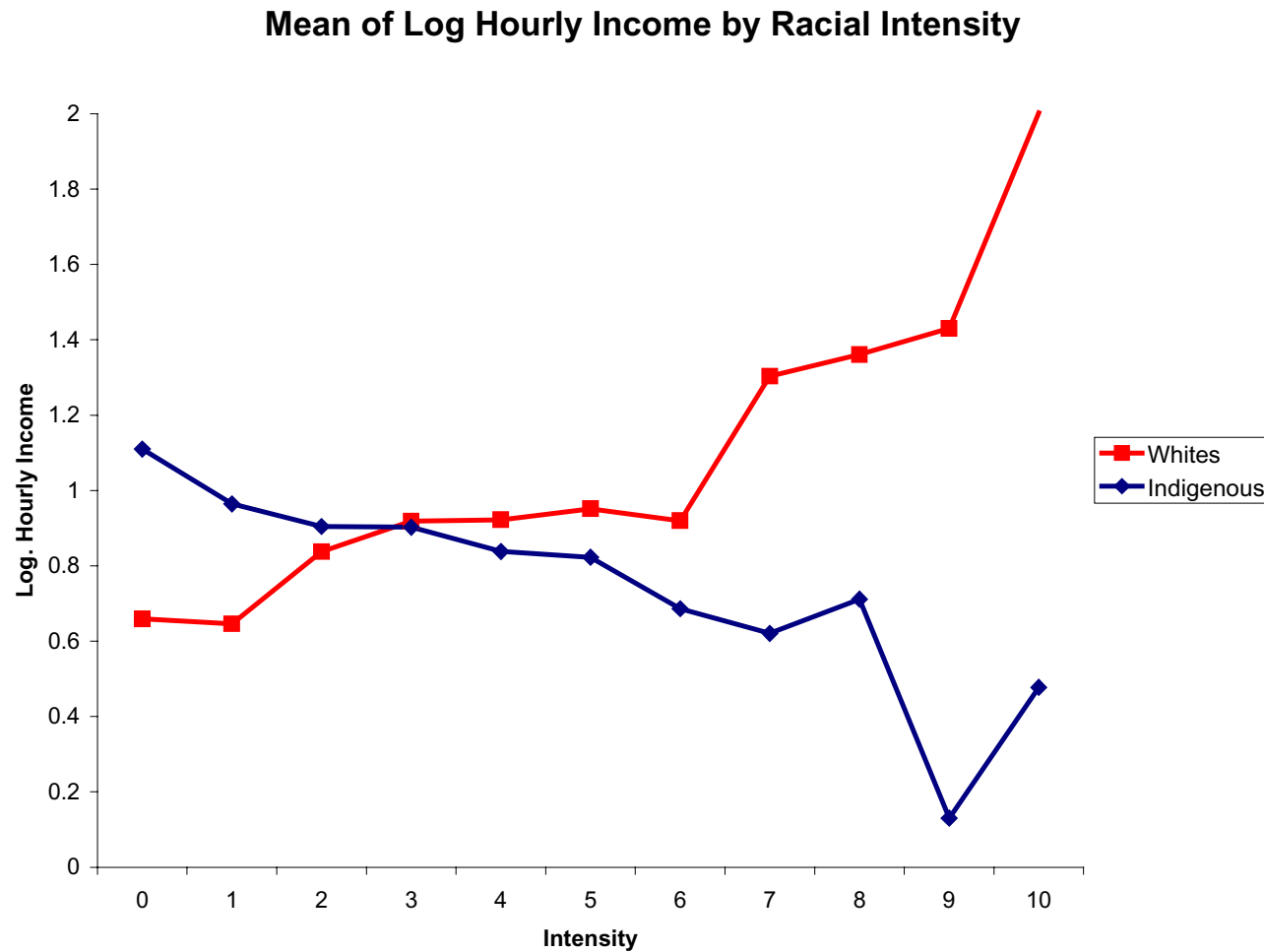
## White Intensity Distribution



## Indigenous Intensity Distribution

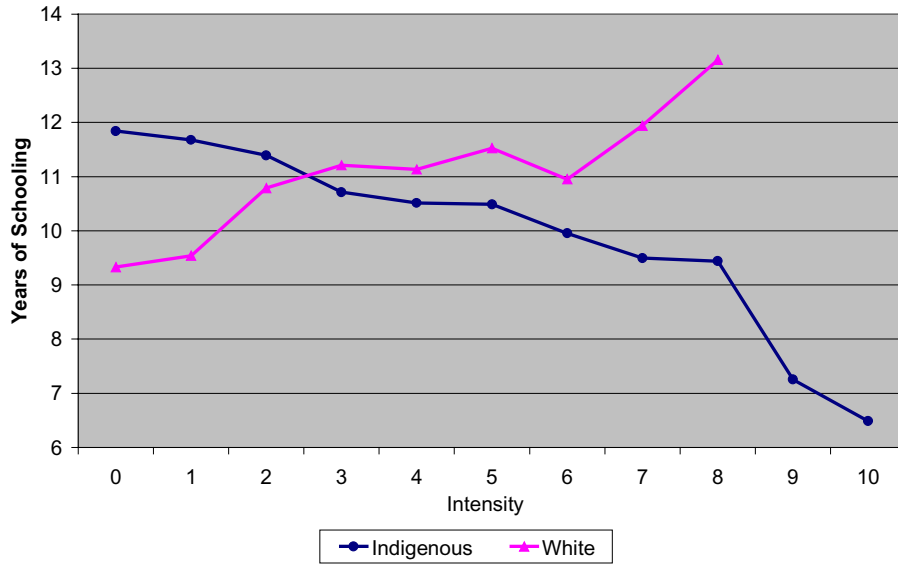


# Racial Intensity and Earnings

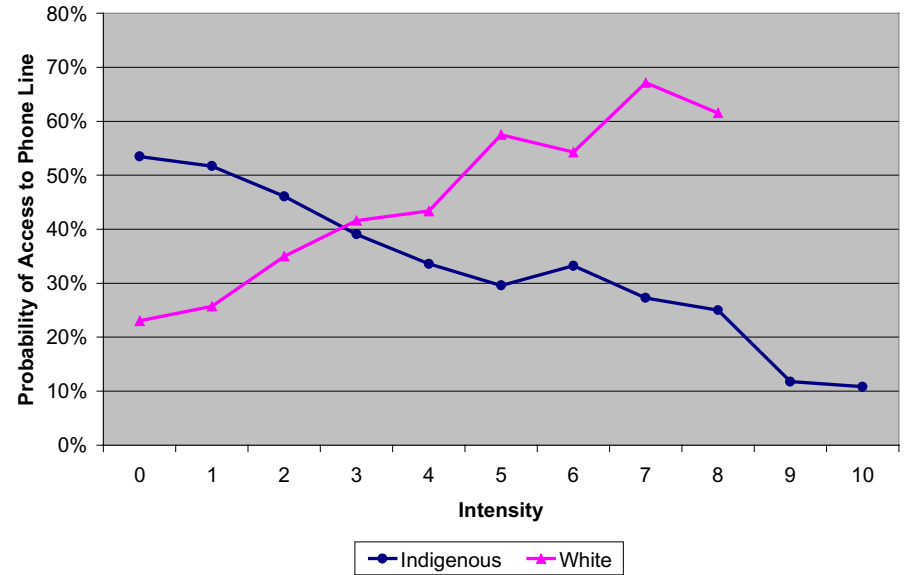


# Individual and Household Characteristics by Racial Intensity

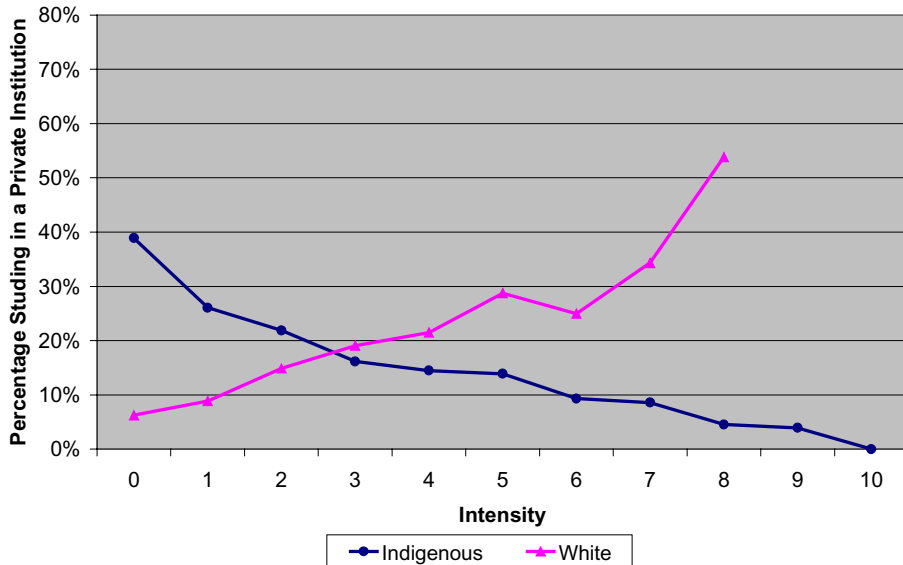
## Years of Schooling by Racial Intensity



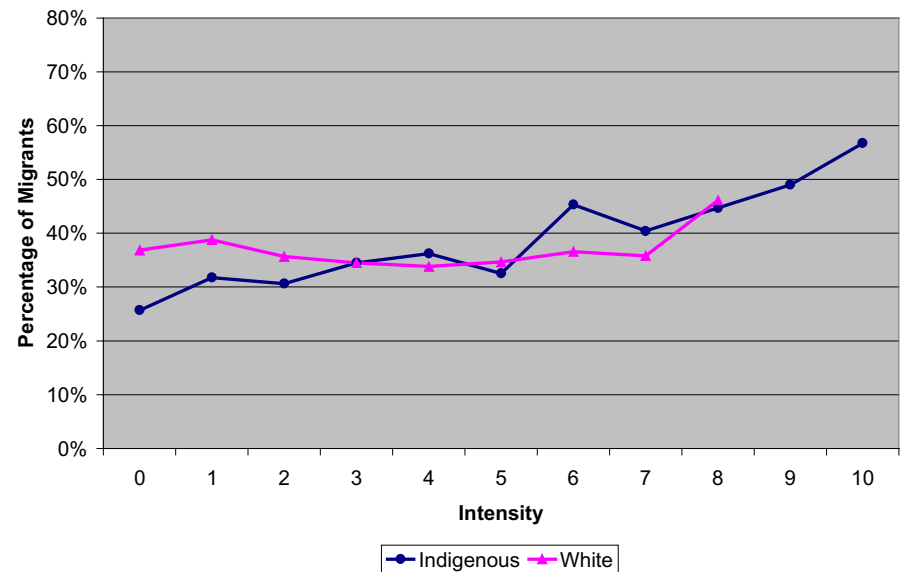
## Access to Phone Line by Racial Intensity



## Studies in a Private Institution by Racial Intensity

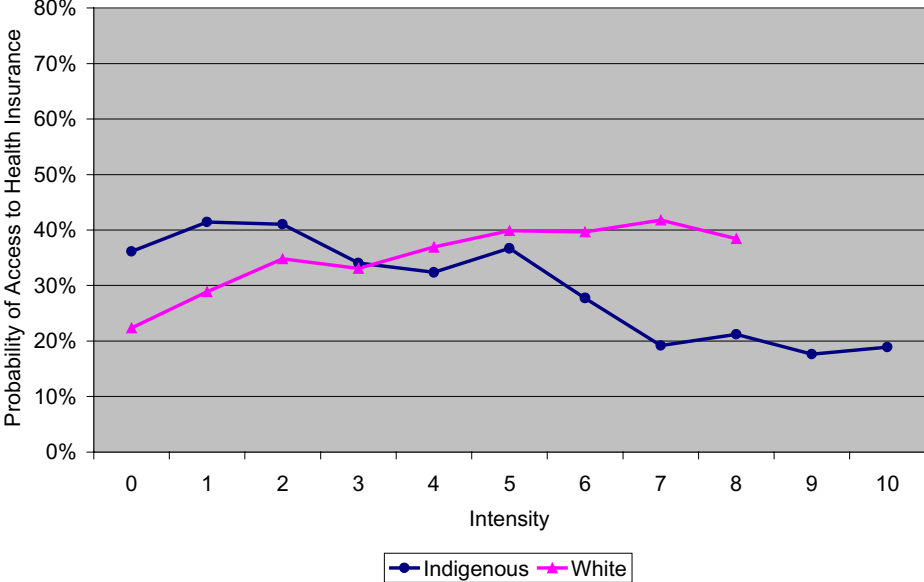


## Migrant Status by Racial Intensity

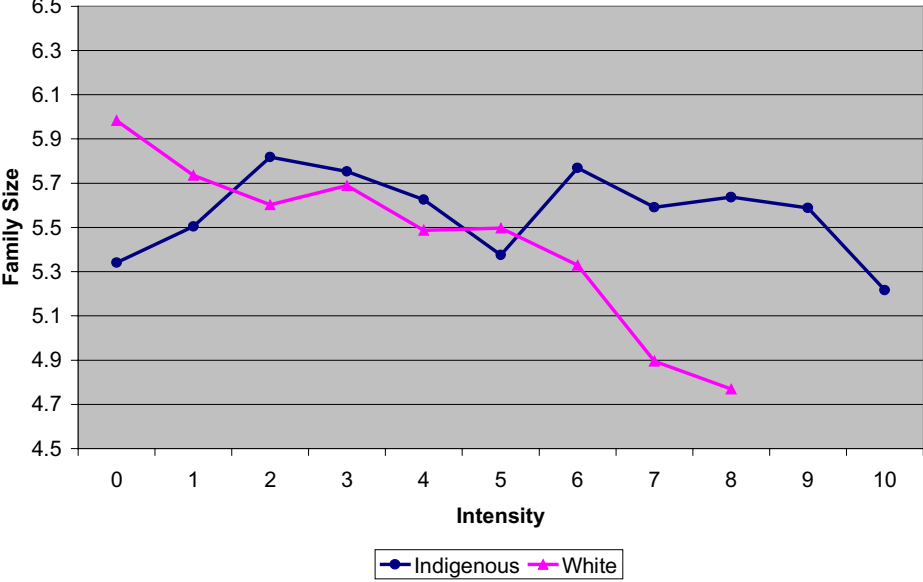


# Individual and Household Characteristics by Racial Intensity

### Access to Health Insurance by Racial Intensity



### Family Size by Racial Intensity



# Individual and Household Characteristics by Religion and Mother Tongue

	<b>Religion</b>		<b>Mother Tongue</b>	
	<b>Catholic</b>	<b>Christian non Catholic</b>	<b>Spanish</b>	<b>Native Language</b>
Years of schooling	10.7	9.2	10.9	9.4
Access to Phone Line (%)	38.6%	22.1%	39.8%	28.3%
Studied in a Private Institution (%)	16.8%	12.8%	19.3%	6.9%
Migrant (%)	35.6%	39.9%	32.9%	45.6%
Access to Health Insurance (%)	34.6%	20.6%	34.8%	27.7%
Family Size	5.6	5.9	5.6	5.8

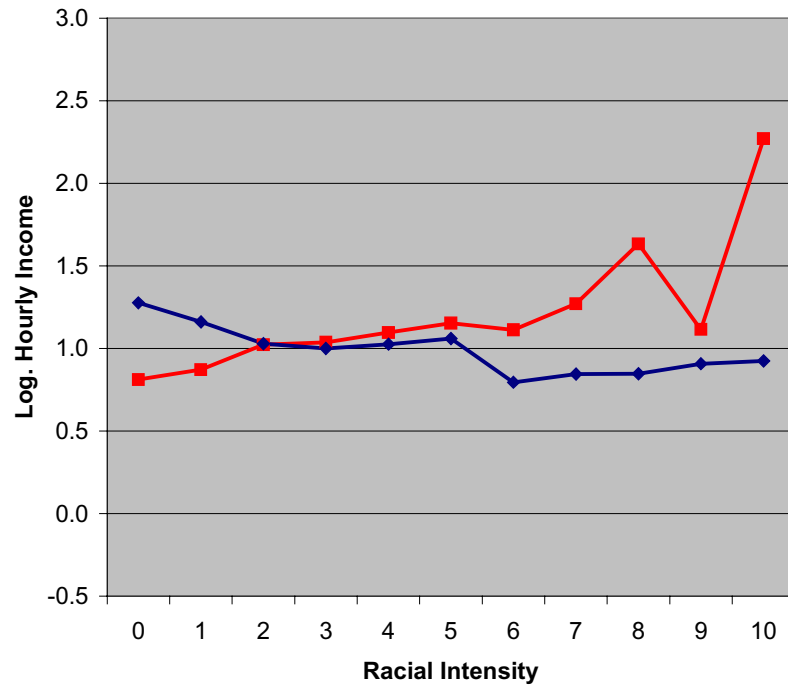
1/ (Including ESSALUD, Private and Others)

Sample: Labor Force without outliers

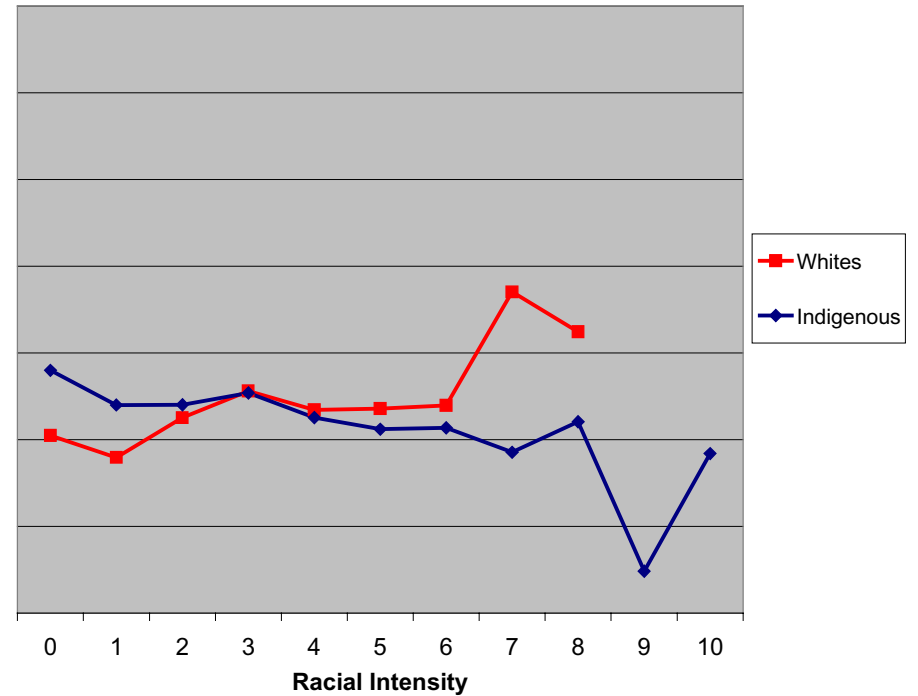
# Wage Earners vs. Self-Employed Income

## Mean of Log Hourly Income by Racial Intensity

Wage Earners



Self-Employed



# Parametric Estimations

- Earnings equations with ethnic variables (religion, mother tongue and polynomial of racial intensities).

$$\ln y = \beta x + \alpha e + \gamma(z) + \varepsilon$$

Where:

In  $y$ : hourly earnings

$x$ : individual characteristics (sex, age, years of education, married), job characteristics (firm size, occupation, economic sector)

$e$ : mother tongue, religion

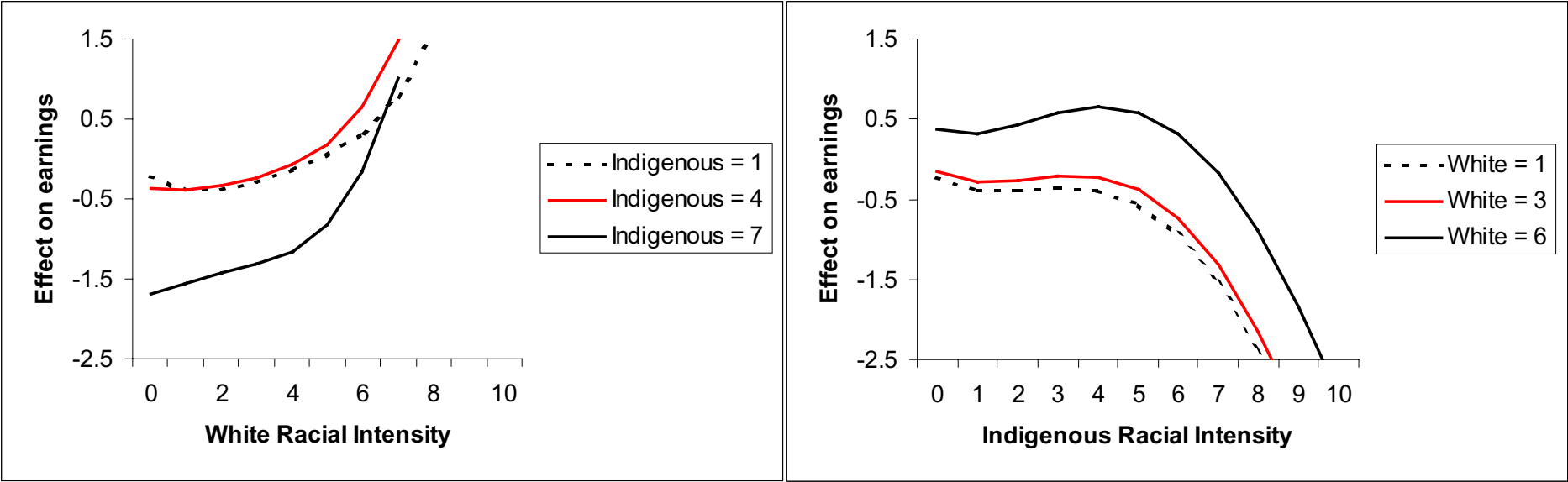
$\gamma(z)$  : polynomial using white and indigenous intensities

# Earnings Equations: Urban Peru

	Wage Earners			Self-Employed		
	(1)	(2)	(3)	(1)	(2)	(3)
Mother tongue native	0.004 (0.04)	0.006 (0.06)	0.032 (0.30)	-0.030 (0.35)	-0.002 (0.02)	0.010 (0.11)
Catholic	-0.157 (0.66)	-0.126 (0.55)	-0.122 (0.54)	0.078 (0.32)	0.113 (0.49)	0.111 (0.47)
Christian non-catholic	-0.188 (0.75)	-0.149 (0.62)	-0.160 (0.67)	0.029 (0.12)	0.068 (0.28)	0.075 (0.30)
No religion	-0.371 (1.11)	-0.340 (1.06)	-0.295 (0.92)	0.362 (1.00)	0.366 (1.05)	0.275 (0.78)
<b>Controls by:</b>						
Individual Characteristics (sex, age, marital status, years of schooling, number of days sick, mothers education)	Yes	Yes	Yes	Yes	Yes	Yes
Labor Characteristics (Occupation, Industry and Firm Size)	No	Yes	Yes	No	Yes	Yes
Polynomial of race intensity	No	No	Yes	No	No	Yes
Observations	1038	1038	1038	1294	1294	1294
Adjusted R-squared	0.22	0.30	0.31	0.16	0.22	0.22

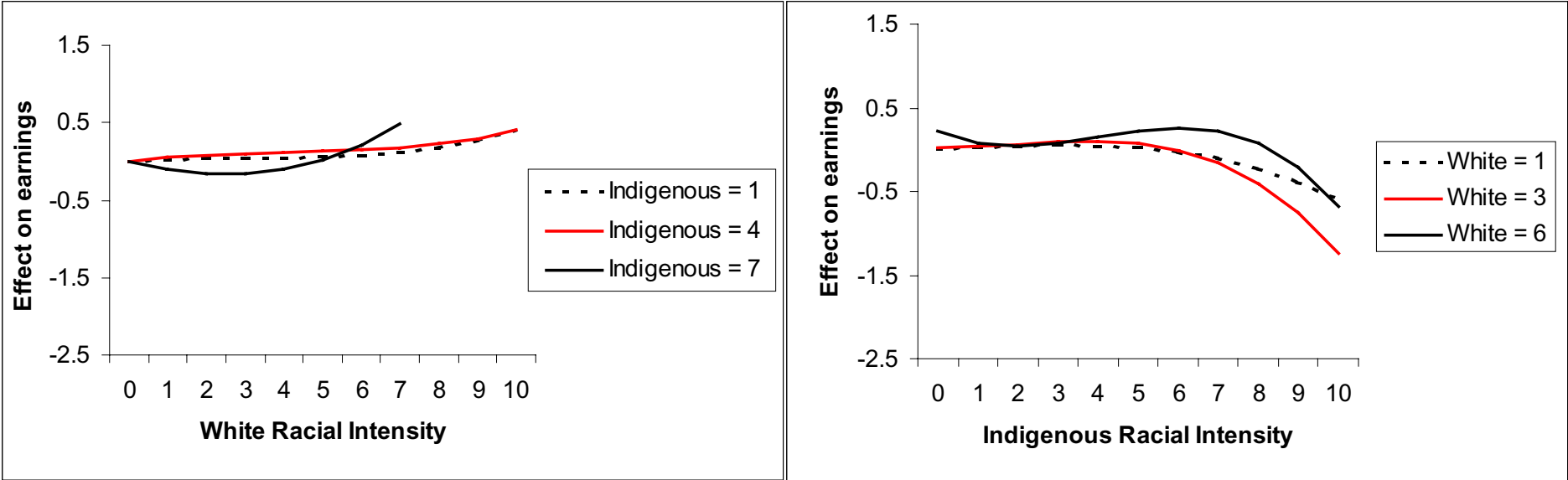
# Racial earnings effects

## Private Wage Earners



# Racial earnings effects

## Self-Employed



# Problems with Parametric Estimations

- We have to impose ex-ante particular polynomial forms to capture racial effects.
- Parametric assumptions not suitable (normality)
- Solution: Semiparametric approach, as in Robinson (1985) approach

# The Semi-Parametric Approach

- Individual and labor-related characteristics are treated as in any regular earnings function
- The racial characteristics are treated in a non-parametrical way
- Quintiles of White and Indigenous race intensity

# The Semi-Parametric Approach

We estimate a semi-parametric log hourly earnings function

$$\ln y = \beta x + \varphi(Z_w, Z_i) + \varepsilon$$

where:

$$\begin{aligned} \beta x &= \text{linear specification of a wage equation} \\ \varphi(Z_w, Z_i) &= \text{non-parametrical estimator for the effect} \\ &\quad \text{of race intensities on earnings} \end{aligned}$$

We use only White and Indigenous race intensity variables.

$Z_w$  and  $Z_i$  are the intensity quintiles in which the individual is situated in the White and Indigenous intensity distribution in the labor force, respectively. The first quintile is the one with the lowest intensity. If individual  $j$  is in the first intensity quintile of the White distribution and in the fourth quintile of the Indigenous distribution,  $Z_w = 1$  and  $Z_i = 4$  respectively. The effect on earnings will be measured by  $\varphi(1,4)$ .

# Estimation procedure

We estimate 
$$\ln y = \beta x + \varphi(Z_w, Z_i) + \varepsilon \quad (1)$$

Applying conditional means we obtain:

$$E[y | Z_w, Z_i] = \beta E[x | Z_w, Z_i] + \varphi(Z_w, Z_i) \quad (2)$$

And

$$y - E[y | Z_w, Z_i] = \beta(x - E[x | Z_w, Z_i]) \quad (3) [(1) - (2)]$$

Let  $b$  be the OLS estimator obtained for equation (3).

We recover  $\varphi(Z_w, Z_i)$  from equation (2):

$$\varphi(Z_w, Z_i) = E[y | Z_w, Z_i] - \beta E[x | Z_w, Z_i] \quad (4)$$

Robinson (85) shows that using this procedure, the linear estimators obtained are root-n consistent and well behaved.

# Construction of Race Intensity Quintiles

- For Whites:

- Primary sorting key: White intensity (ascending)
- Secondary sorting key: Indigenous intensity (descending)

- For Indigenous:

- Primary sorting key: Indigenous intensity (ascending)
- Secondary sorting key: White intensity (descending)

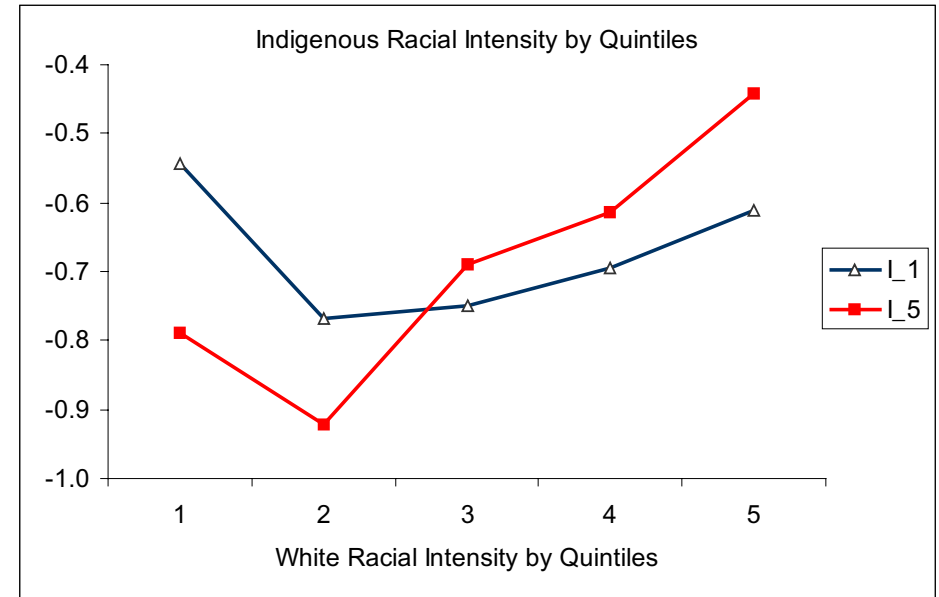
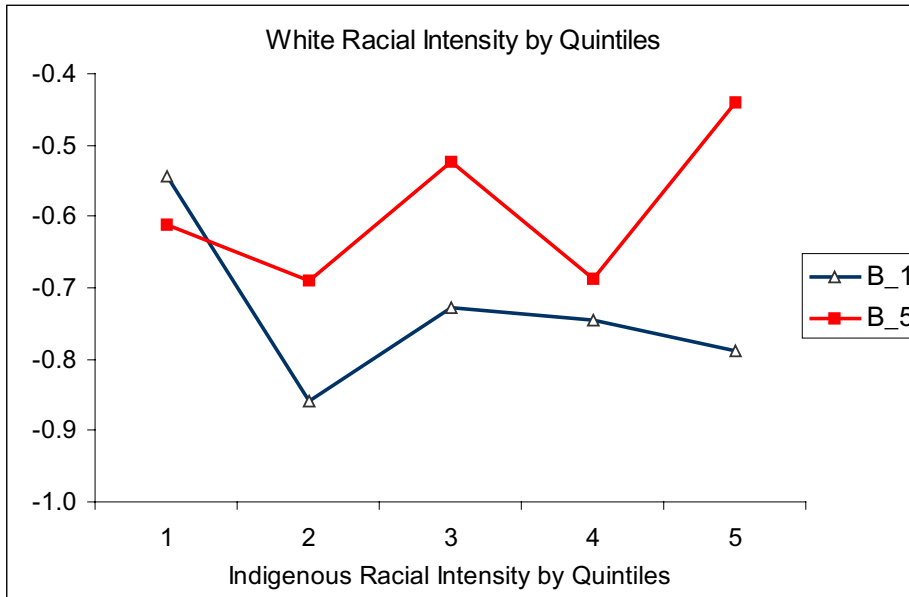
# Non-Linear Estimations: Private Wage Earners

$$\varphi(Z_w = 1, \bullet)$$

$$\varphi(Z_w = 5, \bullet)$$

$$\varphi(\bullet, Z_i = 1)$$

$$\varphi(\bullet, Z_i = 5)$$



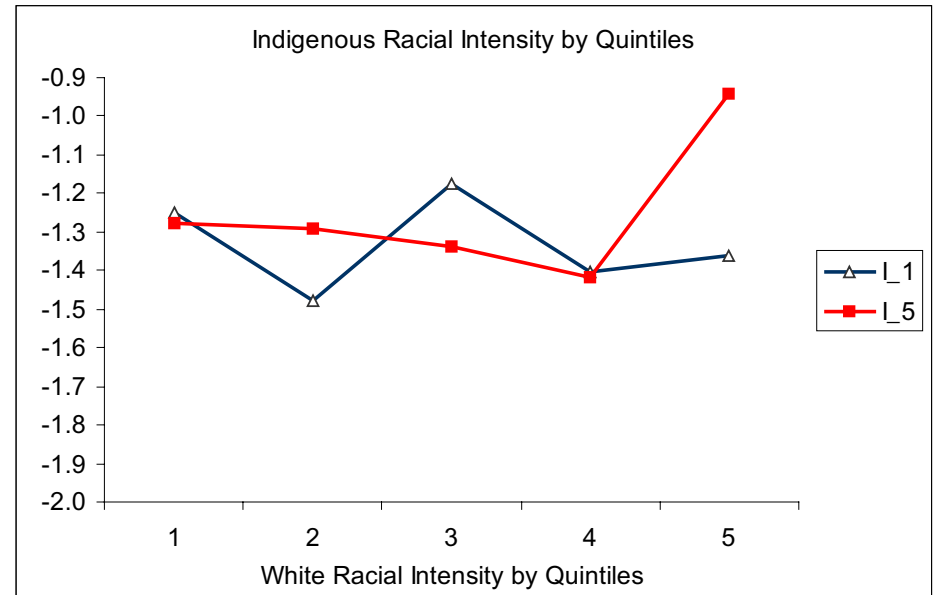
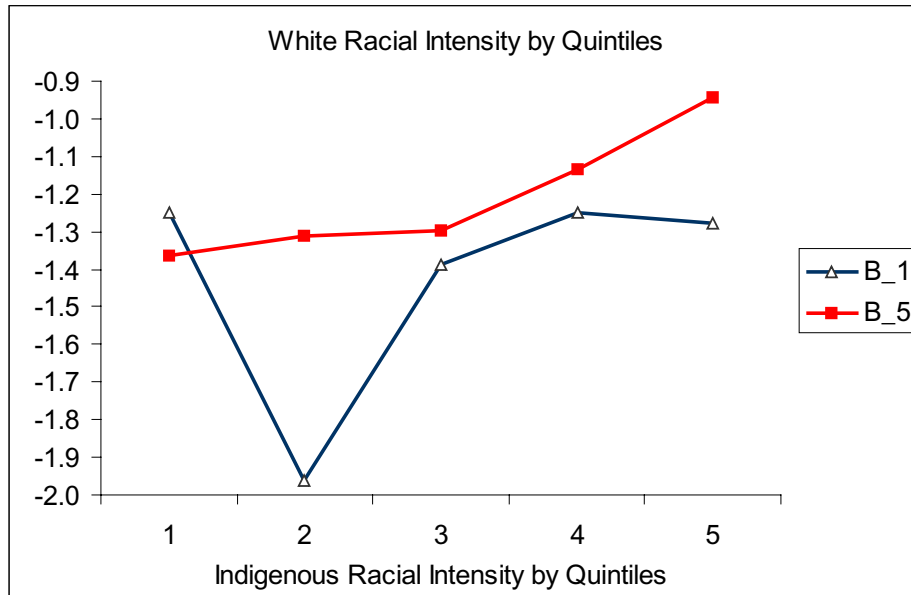
# Non-Linear Estimations: Self-Employed

$$\varphi(Z_w = 1, \bullet)$$

$$\varphi(Z_w = 5, \bullet)$$

$$\varphi(\bullet, Z_i = 1)$$

$$\varphi(\bullet, Z_i = 5)$$



# Linear Estimations

	Wage Earners		Self-Employed	
	(1)	(2)	(1)	(2)
Mother tongue native	0.008 (0.07)	0.023 (0.21)	-0.026 (0.29)	-0.008 (0.09)
Catholic	-0.160 (0.67)	-0.120 (0.52)	0.032 (0.13)	0.087 (0.37)
Christian non-catolic	-0.208 (0.83)	-0.156 (0.64)	-0.006 (0.02)	0.041 (0.17)
No religion	-0.353 (1.06)	-0.289 (0.89)	0.307 (0.84)	0.365 (1.03)
<b>Controls by:</b>				
Individual Characteristics (sex, age, marital status, years of schooling, number of days sick, mothers education)	Yes	Yes	Yes	Yes
Labor Characteristics (Occupation, Industry and Firm Size)	No	Yes	No	Yes
Observations	1038	1038	1294	1294
Adjusted R-squared	0.20	0.26	0.14	0.19

# Earnings Differentials:

## Empirical Distribution

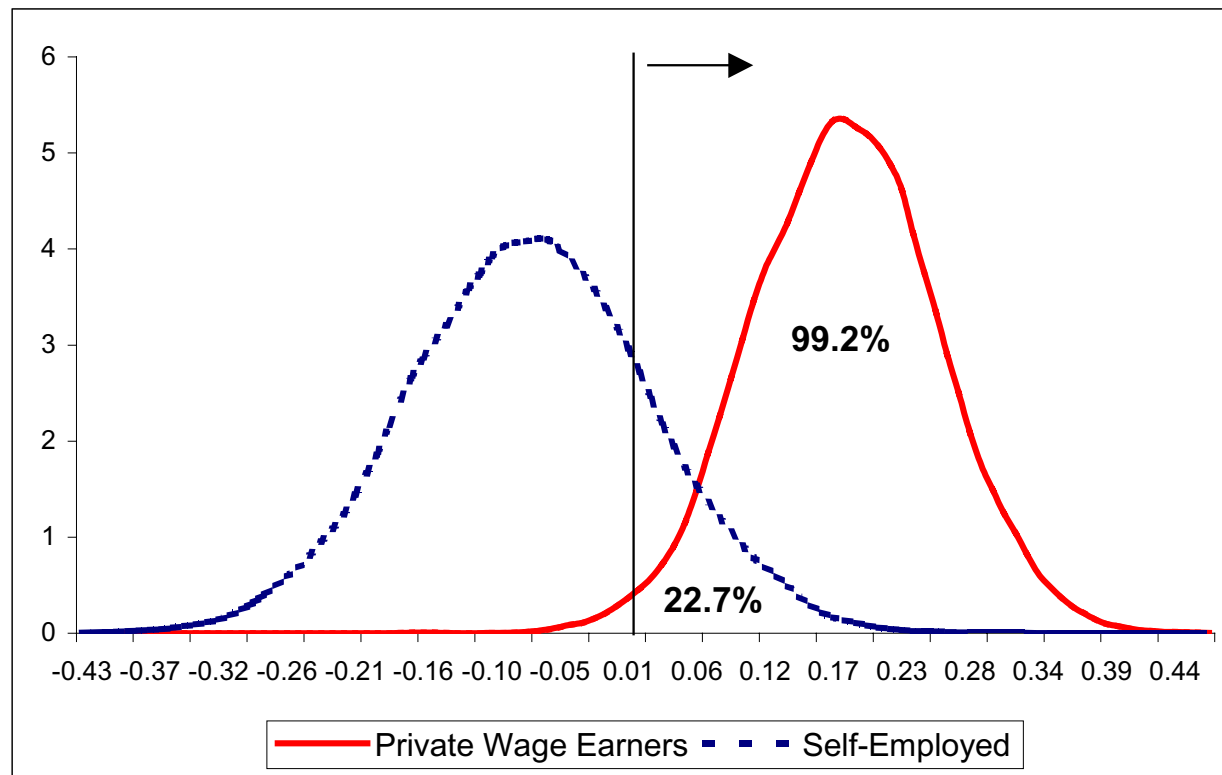
- Re-sampling technique (bootstrap) generating pseudo-samples
- Estimate the racial effects for each pseudo-sample
- Empirical joint distribution of the racial effects
- Empirical distribution for the difference of any pair of effects

# Testing differences using the empirical distributions

- For the non-linear component of our equation we constructed, by means of bootstrapping, an empirical joint distribution for all the possible 25 racial effects.
- We test the difference between the race effects for individuals of different quintile combinations, for instance  $\varphi(5,1) - \varphi(1,5)$ , the predominantly white with respect to the predominantly indigenous.
- For each iteration we computed the difference of both effects,  $\varphi(5,1) - \varphi(1,5)$ , such that we are able to construct an empirical distribution of the difference of these two effects, and, in particular we can find the (empirical) probability of having a positive value for that difference.

# Empirical Distribution for the difference of racial effects on earnings:

$$\varphi(Z_w = 5, Z_i = 1) - \varphi(Z_w = 1, Z_i = 5)$$

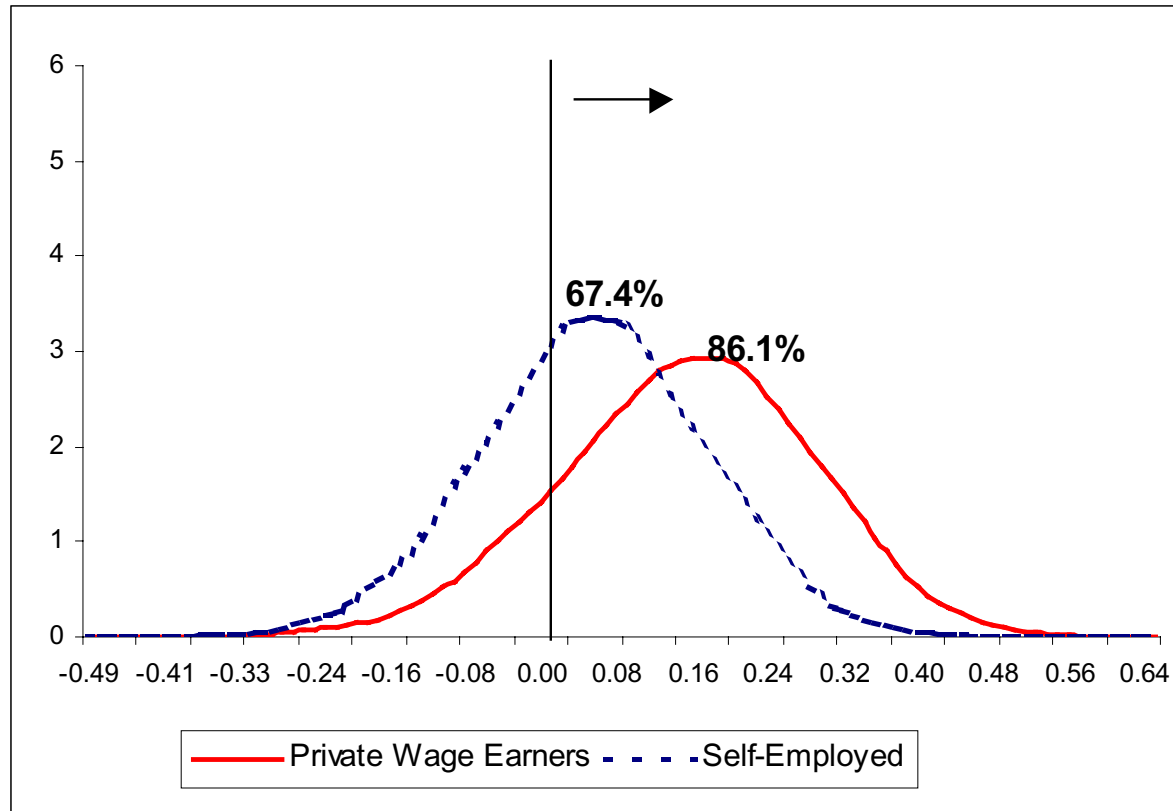


Private Wage Earners: 99.2% of the differences greater than zero

Self-Employed: 22.7% of the differences greater than zero

# Empirical Distribution for the difference of racial effects on earnings:

$$\varphi(Z_w = 5, Z_i = 1) - \varphi(Z_w = 3, Z_i = 3)$$



Private Wage Earners: 86.1% of the differences greater than zero

Self-Employed: 67.4% of the differences greater than zero

# Occupational Segregation

- Duncan Index for Occupational Segregation by Racial Characteristics

$$I_D = \frac{1}{2} \sum_{i=1}^n \left| \frac{W_i}{W} - \frac{I_i}{I} \right|$$

- Benchmark:  
Duncan Index for Occupational Segregation by Gender = 0,2586

# Decomposition of Earnings Gap

- Wage Gap: Whites, Indigenous, “Mestizos”
- Benchmark: Wage Gap by gender = 0,1569
- Results sensibility to “cut-off rules”

$$I \geq x \wedge W < x \Rightarrow \textit{Indigenous}$$

$$W \geq x \wedge I < x \Rightarrow \textit{White}$$

- “Mestizos” are defined as the complement  
“x” : cut-off point (it may take any value between 0 and 10, but in order to get reasonable estimates we consider only values between 3 and 5)
- Characteristics vs. Returns

$$\textit{WageGap} = \beta_i(X_i - X_j) + (\beta_i - \beta_j)X_j$$

Where  $i, j$  are subscripts for white, indigenous and mestizo, and  $i \neq j$

# Wage Gap Decomposition

## Wage Gaps comparison between Whites, Mestizos and Indigenous

	Cut-Off Criteria 5		
	White vs Indigenous	White vs Mestizos	Mestizos vs Indigenous
<b>WAGE GAP</b>	0.412	0.207	0.204
<b>Differences in Characteristics</b>	0.322	0.134	0.131
Education and Experience	0.157	0.058	0.071
Personal and Family Characteristics	0.060	0.025	-0.003
Labor Market Variables <sup>1/</sup>	0.106	0.051	0.064
<b>Differences in Returns</b>	0.089	0.074	0.073

1/ Occupation, Industry and Firm Size

# Wage Gap Decomposition

## Wage Gaps comparison between Whites, Mestizos and Indigenous

	Cut-Off Criteria 4		
	White vs Indigenous	White vs Mestizos	Mestizos vs Indigenous
<b>WAGE GAP</b>	0.347	0.136	0.211
<b>Differences in Characteristics</b>	0.252	0.114	0.124
Education and Experience	0.137	0.048	0.062
Personal and Family Characteristics	0.055	0.019	0.001
Labor Market Variables <sup>1/</sup>	0.060	0.047	0.061
<b>Differences in Returns</b>	0.095	0.023	0.087

1/ Occupation, Industry and Firm Size

# Wage Gap Decomposition

## Wage Gaps comparison between Whites, Mestizos and Indigenous

	Cut-Off Criteria 3		
	White vs Indigenous	White vs Mestizos	Mestizos vs Indigenous
<b>WAGE GAP</b>	0.356	0.162	0.194
<b>Differences in Characteristics</b>	0.303	0.164	0.111
Education and Experience	0.154	0.075	0.062
Personal and Family Characteristics	0.042	0.035	0.003
Labor Market Variables <sup>1/</sup>	0.107	0.054	0.046
<b>Differences in Returns</b>	0.053	-0.002	0.083

1/ Occupation, Industry and Firm Size

# Duncan Index

---

	Duncan Index
<b>Criteria 5</b>	
White vs Ind+Blacks	0.292
White vs Mestizos	0.254
Mestizos vs Ind+Blacks	0.157
<b>Criteria 4</b>	
White vs Ind+Blacks	0.265
White vs Mestizos	0.214
Mestizos vs Ind+Blacks	0.155
<b>Criteria 3</b>	
White vs Ind+Blacks	0.293
White vs Mestizos	0.248
Mestizos vs Ind+Blacks	0.150

---